University of Warwick Department of Economics

Corruption & Growth

University ID: 0400544[ψ]

EC331 Research in Applied Economics

April 2007

Abstract

This paper re-examines the direct and indirect effects of corruption on economic growth in a cross-country growth framework covering 129 countries over the period 1970-2000, using corruption perceptions' as a proxy for actual corruption. While two-stage least-squares estimation would be useful to account for the endogeneity of corruption, all of the instruments are found to be weak and quantitative results are shown to be unreliable since the effects of reverse causality are insufficiently accounted for.

[ψ] **Acknowledgement:** The author is grateful to Professor Nicholas Crafts for his supervision and helpful discussion prior to the write-up of this research paper. All possible errors and omissions are the sole responsibility of the author.

Contents

'Optima corrupta pessima' – The best things, once corrupted, become the worst

Latin Proverb

Section I

Introduction

The general consensus that has emerged from the vast recent literature on corruption, defined by the World Bank as the abuse of public power for private benefit, as a serious impediment to economic growth was underlined when the United Nations '*Convention Against Corruption*' came into force in December 2005. However, according to the World Bank, poverty itself might cause corruption because poor countries cannot devote sufficient resources to setting up and enforcing an effective legal framework, or because people in need are more likely to abandon their moral principles. Indeed, if high levels of corruption are associated with low levels of GDP, it is not immediately clear whether this is because high corruption will cause low growth or low growth will cause high corruption.

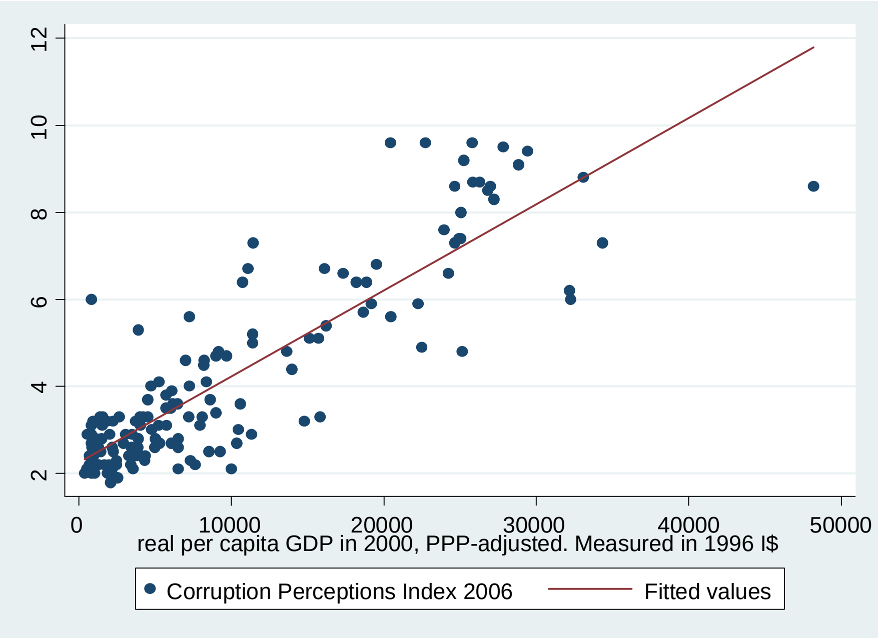

Figure 1 Scatter plot of Transparency International's Corruption Perceptions' Index 2006 with real GDP per capita in 2000, PPP adjusted.

Figure 1 provides an illustration of the strong correlation that exists between corruption and economic development for a cross-section of 129 countries. The technique that has become standard to correct for endogeneity is the use of instrumental variables or two-stage least-squares. This requires the identification of an instrument that is highly correlated with corruption and uncorrelated with the error term in the growth regression. However, the instruments commonly used for corruption, ethnolinguistic fractionalization and legal origins, have been found to be weak. Hence, the issue of causality is still pressing.

In a pure cross-section of 129 countries over the period between 1970 and 2000, this study first investigates the feasibility of using other datasets that have emerged recently as possible instruments for corruption. Second, the direct and indirect transmission channels through which corruption affects growth levels, and the relative contribution of various channels to the overall effect of corruption on growth, are explored. Specifically, the focus is on the effect of corruption on investment, schooling, trade policy, political instability and natural resources using variables most commonly used in the growth literature.

Section II
Literature Review

The theoretical argument that corruption would tend to lower economic growth through investment as the main channel variable was set out by authors such as Shleifer & Vishny (1993) and further reinforced by influential empirical work by Mauro (1995) using a two-equation estimation procedure. Simultaneously, the emergence of extensive cross-country data provided by organizations such as Transparency International (TI) and the World Bank encouraged further work on the importance of institutions in general, and corruption in particular. The apparent finding that economic growth will be faster in the absence of corruption led to a keen interest in the main causes of corruption in the fight against poverty. However, extensive econometric studies on the causes of corruption such as Treisman (2000), find that lower economic development, usually measured as real GDP per capita, to be a significant cause of corruption. Low income results in poor institutional settings which, in turn, create incentives for civil servants to collect bribes. Thus, there is theoretical and empirical evidence that causality works both ways with regards to corruption and growth.

One of the founders of the political-economic literature on corruption, Rose-Ackerman (1999) described corruption as an epidemic with economic, cultural and political nuances. Hence, it should not come as a great surprise that subsequent cross-sectional empirical studies have found many different variables to be correlated to corruption. Mauro (1995) controlled for the effect of endogeneity by the use of an index for ethnolinguistic fractionalization (ELF), as an instrumental variable (IV) for corruption. In addition, La Porta, Lopez-de-Silanes, Shleifer & Vishny (1999) argued that both legal origin and ethnolinguistic fractionalization explain the quality of government and added further weight to the increasingly-popular belief that institutions are important determinants of economic and financial outcomes. Pellegrini & Gerlagh (2004) uses legal origins as an instrument and further finds that investment and trade openness are the main channels through which corruption affects long-run growth.

However, in recent studies the instrumental variables used to control for endogeneity by two-stage least squares (2SLS) methods in the literature on corruption such as the ELF index and legal origins have been criticized and are shown to be weakly identified by Shaw, Katsaiti & Jurgilas (2006). They argue that neither serves as valid instruments for corruption and hence, the results of previous studies need to be interpreted with caution. Weak instruments that are not well-correlated with corruption can lead IV estimates to be even more biased than standard OLS estimates. One agenda that stems from this conclusion is the search for new variables that serve as better instruments.

On that front, two recent papers are highlighted. The first, by Acemoglu & Johnson (2005) exploited differences in the historical experiences of former European colonies and found that contracting institutions, instrumented by legal origins, appear to matter only for the form of financial intermediation whereas property rights institutions, instrumented by colonization strategy (settler mortality and population density), have a first-order effect on long-run economic growth, investment, and financial development. This result is interesting since the most commonly used definition of corruption, the abuse of public power for private benefit, bears more resemblance to the definition of property rights institutions, those which protect citizens against expropriation by the government and powerful elites, rather than that of contracting institutions, those which enable private contracts between citizens. Thus, colonization strategy can be viewed as better instruments for corruption. Additionally, the theoretical argument that the Europeans' colonization strategy was based on mortality rates means it is also possible to exploit differences in past malaria rates.

The second, by Alesina, Devleeschauwer, Easterly, Kurlat & Wacziarg (2003), constructed more comprehensive measures of ethnic, linguistic and religious fractionalization and finds that ethnic and linguistic fractionalization in particular are associated with negative outcomes in terms of quality of government, which could be interpreted as a broader measure of corruption. Furthermore, Roeder (2001) provides alternative measures of ethnolinguistic fractionalization. To the best of the authors' knowledge, none of these indices have been used as instruments in the corruption literature.

In the authors' own opinion, taking a sample of less than a third of the total number of countries in this world is insufficient to make generalizations about the rest of the sample. Shaw et al (2006) emphasise the importance of reporting the first-stage F-stat (or t-stat) on the instrument(s) to avoid making erroneous inference. Since the seminal paper by Mauro (1995), a vast body of research has focused on 2SLS methods, thus revealing the dangers of estimation and inference under weak identification, validating this revisiting of the empirical literature.

The literature on corruption often links closely to the natural resource curse thesis, the tendency for countries with an abundance of natural resources to have lower economic growth. Sachs & Warner (1995, 2001) find that evidence for the natural resource curse works through the Dutch disease effect, controlling for variables such as geography and climate. The Dutch disease effect is said to occur as the discovery of a natural resource raises the value of a nation's currency, making manufactured goods less competitive with other nations, increasing imports and decreasing exports. However, more recent studies by Leite & Weidmann (1999) and Sala-i-Martin (2003) argue that natural resource abundance creates opportunities for rent-seeking behavior and is an important factor in determining a country’s level of corruption.

Authors such as Acemoglu, Johnson & Robinson (2001, 2002) showed that institutions are persistent over time and are fundamental determinants of economic growth. Pellegrini & Gerlagh (2004) note that this implies corruption levels tend to persist over time. Meanwhile, Sachs & Warner (2001) noted that although data is limited, the stagnation in economic growth of resource-abundant countries began in the early 1970s. Hence, whilst corruption tends to persist over time, the natural resource discoveries that have a negative impact on growth are fairly recent. The presence of corrupt institutions in a country means the discovery of natural resources would exert a negative impact on growth rather than a positive impact. This study includes natural resources as a transmission mechanism through which corruption negatively affects growth.

The present paper directly extends the analysis of Pellegrini & Gerlagh (2004) who measured the contribution of the direct and indirect effects of corruption on economic growth, using a two-equation estimation procedure validated only by the Hausman Test, with the precautionary reminder of Shaw et al (2006) in mind. A larger sample from more recent data from

Transparency International serves as a further check on the robustness of their findings. This study also further includes the natural resource curse as a channel through which corruption negatively affects growth.

Section III
Methodology

In this section, the methodology employed will be explained in brief. The first part describes the basic cross-country growth regression, followed by two-stage least-squares technique to possibly correct for endogeneity. The third part summarises the contributions of the transmission channels, and finally, the long-term effects are calculated to analyse the pervasiveness of corruption on growth.

III.1 Cross-Country Growth Regressions

$$G^i = \alpha_0 + \alpha_1 \ln(Y_0^i) + \alpha_2 C^i + \alpha_3 Z^i + \varepsilon^i \quad (1)$$

Initially, a basic growth regression is estimated to quantify the effect of corruption on economic growth without, and then with, other independent variables. This is based on the common regression equation (1) where the dependent variable G denotes the long-run GDP growth rate. As common in the growth literature, the (natural logarithm of) the level of initial income, $\ln(Y_0^i)$ is included as an independent variable and is expected to have a negative coefficient, $\alpha_1 < 0$, according to the conditional convergence hypothesis, which states that the growth rate of income is expected to be negatively associated with the level of income at the beginning of the period. The coefficient α_2 of the second independent variable is corruption C^i and is the main subject of the analysis. The vector Z^i is the transmission channels that are possibly linked with corruption namely investment, trade openness, schooling, political instability and natural resources using variables common in the growth literature. These results serve as a robustness check for Pellegrini & Gerlagh (2004).

III.2 Two-Stage Least Squares

As stated earlier, there exists theoretical and empirical evidence that causality works both ways with regards to corruption and growth. The technique that has become standard to correct for endogeneity is the use of instrumental variables or two-stage least-squares. The endogeneity of corruption in the growth regression is checked formally by the Hausman test, to ensure 2SLS is necessary (i.e. the null hypothesis that the regressor is exogenous is rejected).

Stage 1:

$$Corruption^{i} = \alpha_2 + \theta GrowthVariables^{i} + \lambda Instruments^{i} + \varphi ChannelVariables^{i} + \eta^{i}$$

Stage 2:

$$GDPGrowth^{i} = \alpha_1 + \beta GrowthVariables^{i} + \gamma Corruption^{i} + \delta ChannelVariables^{i} + \varepsilon^{i}$$

Furthermore, 2SLS requires the identification of an instrument that is highly correlated with corruption and uncorrelated with the error term in the growth regression. The first requirement[1] can be tested using a *t* or *F* test in the reduced form regression (Stage 1). The second requirement[2] can be tested if there is more than one[3] IV by Sargan's test for overidentifying restrictions (i.e. the null hypothesis that instruments are uncorrelated with the error term is not rejected – a rejection casts doubt on the validity of the instruments).

[1] This requirement was not reported in Pellegrini & Gerlagh (2004) and it could be that legal origins and ethnolinguistic fractionalization are weak instruments for corruption.

[2] This requirement was not explored in the analysis by Pellegrini & Gerlagh (2004) as they only use one instrument for corruption.

[3] Under standard 2SLS assumptions, the addition of one more instruments to the list improves the asymptotic efficiency of the 2SLS, but the addition of too many instruments – increasing the number of overidentifying restrictions – can cause severe biases in 2SLS. The instruments must be exogenous to growth; otherwise 2SLS will not be consistent.

One feature of the IV estimator is that if corruption is endogenous to growth – so that instrumental variables estimation is actually needed – it is essentially never unbiased. Hence, the IV estimator can have substantial bias in small samples, and large samples are preferred. The potential drawback with using IV is that the estimates can be inconsistent and have large standard errors, if the instrument is only weakly correlated with corruption. In the absence of weak instruments, OLS estimation can be preferred to IV. Correlation among regressors can lead to large standard errors for the OLS estimates, but can be even more serious with 2SLS. Shaw et al (2006) noted that the ELF index and legal origins are weak instruments for corruption. This validates a re-visit of the literature, by exploring other variables as instruments for corruption that have received less attention.

III.3 Direct and Indirect Effects of Corruption

The dependence of the variables in the Z^i vector on corruption is estimated according to the following equation:

$$Z^i = \beta_0 + \beta_1 \ln(Y_0^i) + \beta_2 C^i + \mu^i \quad (2)$$

where β_0, β_1 and β_2 are five-dimensional vectors of the coefficients; β_1 describes the effect of income on the vector of dependent variables Z^i and μ^i is the vector of residuals. The direct and indirect effects of corruption on economic growth can be singled out by the formal substitution of equation (2) into (1) :

$$G^i = (\alpha_0 + \alpha_3\beta_0) + (\alpha_1 + \alpha_3\beta_1)\ln(Y_0^i) + (\alpha_2 + \alpha_3\beta_2)C^i + \alpha_3\mu^i + \varepsilon^i \quad (3)$$

where α_2 is the direct effect of corruption on economic growth and $\alpha_3\beta_2$ captures the summed indirect effects of corruption on economic growth, and μ^i is the residuals of equation (2). The contribution of the direct effect relative to the total effect is $\alpha_2 / (\alpha_2 + \alpha_3 \beta_2)$.

III.4 Long-Term Effect of Corruption on the Transmission Channels

Finally, the long-term effect of corruption on the transmission variables is calculated. There are two approaches highlighted by Pellegrini & Gerlagh (2004) through which long-term effects can be estimated and comparing the result from both approaches provides a robustness check on the findings. The first approach is:

$$\Delta Z_{\infty} / C = \beta_2 + \left(- (\alpha_2 + \alpha_3 \beta_2) / (\alpha_1 + \alpha_3 \beta_1) \right) \beta_1 \quad (4)$$

where β_2 is the direct effect of corruption on the transmission variables (from equation (2)) and measures the effect of corruption on the transmission variables abstracting from income effects. The second term $\left(- (\alpha_2 + \alpha_3 \beta_2) \beta_1 / (\alpha_1 + \alpha_3 \beta_1) \right)$, multiplies the long-term income effect[4] of corruption (based on coefficients as expressed in equation (3)) multiplied by the effect of income on the transmission variables (from equation (2)) and can be interpreted as the long-term effect of corruption on the transmission variables operating through the income variable.

As mentioned earlier, institutions are persistent over time and it suggests that they are the main determinants of the long-term economic performance. As a check the second approach directly estimates the dependence of the transmission channel variables on the corruption variable, omitting the initial income as an explanatory variable. The specification of equation (2) becomes

$$Z^i = \gamma_0 + \gamma_1 C^i + \nu^i \quad (5)$$

[4] See Pellegrini & Gerlagh (2004) Appendix 1 for derivation of long-term income effects.

where the variable $\ln(Y_0^i)$ has been omitted as an independent variable. The vector of coefficients γ_1 describes the effect of corruption on the vector of dependent variables Z^i and v^i is the vector of residuals. If corruption is indeed persistent over time, then the coefficients should be of similar magnitude.

The estimates of the coefficients in equation (4) are expected to be slightly larger in absolute value than those in equation (5) due to the fact that equation (4) implicitly assumes the level of corruption to be constant over time, and thus gives too much weight to the future effect of present corruption levels.

Section IV
Data Description

This study takes a purely cross-sectional approach, using data from 129 countries between the period 1970 and 2000. The countries were chosen where data on corruption and growth were both available, but the data set is limited in that some variables were not available for this period exactly, and there are missing values for certain countries.[5] The data were gathered mainly from the Penn World Tables 6.2 and the Harvard Centre for International Development.

IV.1 Growth Variables

The main dependent variable is the long-run growth rate of output (*growth7000*) and is measured as the average annual growth rate of GDP 1970-2000: $G^i = (1/T)\ln\left(Y_T^i / Y_0^i\right)$ where $t_0 = 1970$ and $t_T = 2000$. Consistent with the growth literature, the natural log of real GDP per capita in 1970 (*lgdp70*),

[5] The countries in the sample are listed in Appendix A, while full variable definitions are listed in Appendix B. A table of means, standard deviations, minima and maxima for each variable can be found in Appendix C.

measured in 1996 international dollars, is included as a parameter which captures conditional convergence.

IV.2 Corruption

The main independent variable of interest is corruption, defined as the abuse of public power for private benefit. A measure of corruption perceptions' (*corrupt*) is commonly used as a proxy to actual corruption[6], which is difficult to measure due to its illegal nature. Furthermore, comparisons cannot be made since laws, and its enforcement, differ significantly between countries. The data is based on the Corruption Perceptions' Index for 2006, published by Transparency International. Despite a changing sample and methodology, the ratings are highly correlated over the years which reflect the pervasiveness of corruption. Another feature of the CPI indices is that the standard deviation of countries within each index has fallen over the years which implies an increasingly common perception amongst pollsters and hence, a more accurate measure of corruption[7].

IV.3 Transmission Channels

The transmission channels through which corruption negatively affects growth explored in this paper are investment, trade policy, schooling, political instability and natural resources, based on common regression variables in the growth literature[8].

The investment variable (*invest7000*) represents the percentage of gross investment (public and private) on GDP in the period 1970-2000 from Penn

[6] As Mauro (1995) argued, the advantage of using subjective data on corruption is that it is investor's perceptions of corruption that determines investment which, in turn, determines growth. Furthermore, a successful IV strategy would correct not only for reverse causality & omitted variable bias, but also differential measurement error.

[7] The author notes that there might be endogeneity problems with taking the 2006 CPI index in explaining growth rates over 1970-2000, but the decision is taken due to persistence of corruption and the fact that recent indices are more accurate and increase the sample size.

[8] For example, Levine & Renelt (1992), Sachs & Warner (1995) and Sala-i-Martin (1997)

World Tables. The variable enters the growth regression as the period average and is assumed exogenous as the attempt to instrument the investment ratio with the initial period value of the price level of investment goods was unsuccessful. The trade openness variable[9] (*openness7092*) is defined as the share of years the country has been open in 1970-1992 by Sachs & Warner (1995). According to Pellegrini & Gerlagh (2004), this variable approximately describes the tendency for corrupted societies to raise trade barriers by regulation, thus creating a potential source of influence and income for policymakers, and by custom bureaucracy, a potential source of bribe income for custom officials. The schooling variable (*sec70*) is the gross secondary school enrolment[10] in 1970, as an approximation for the investments in human capital. The variable for political instability (*pinstab7085*) measures the averaged sum of revolutions and the number of assassinations per million people per year in the year 1970-1985. Finally, the measure for resource abundance is the share of total exports of natural resources in GNP in 1970 (*natresgnp70*).

The variable coefficients for investment, trade openness, and schooling are expected to be positively correlated with growth and negatively correlated with corruption, whereas the variable coefficients for political instability and natural resources are expected to be negatively correlated with growth and positively correlated with corruption.

IV.3 Instrumental Variables

The instruments explored are legal origins (*legor_uk*), colonial strategy based on the population density in 1500 and settler mortality rates (*log_pop_dens_1500*, *log_sett_mort*), past malaria rates (*mal_46a*, *mal_46p*) as

[9] On this point, it is important to note that Rodriguez and Rodrik (1999) find that the use of other trade variables may lead to varied results, but for the sake of comparison, this dataset is used.

[10] Pellegrini and Gerlagh (2004) used a different measure of schooling: the average number of schooling years in the population over the age of 25 in 1975. However, this is replaced by secondary school enrollment since it is more widely used in the literature.

well as alternative measures for ethnolinguistic fractionalization (*ethnic, language, elf61* & *elf85*).

Section V
Results and Analysis

V.1 Cross-Country Growth Regressions

Dependent variable: *growth7000*	(1)	(2)	(3)
lgdp70	-0.89***	-0.96***	-0.11***
	(0.19)	(0.27)	(0.29)
invest7000		0.06**	0.06*
		(0.03)	(0.03)
sec70		0.02	0.01
		(0.02)	(0.02)
openness7092		2.56***	2.29***
		(0.51)	(0.59)
pinstab7085		-0.29	0.025
		(1.29)	(1.34)
natres70		-4.21**	-3.97**
		(1.88)	(1.91)
corrupt	**-0.64******		**-0.13**
	(0.09)		**(0.15)**
constant	12.45	7.59	9.18
n	129	83	83
Adjusted R^2	0.27	0.42	0.42

Table 1 Growth regressions based on equation (1)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

Regression (1) in *Table 1* presents the results of estimating equation (1) with the inclusion of only the initial income level and corruption as independent variables. The coefficients have the predicted signs and are statistically significant at 1%, which reflects the strong correlation between corruption and growth. Regression (2) in *Table 1* presents the results of estimating equation (1) with the inclusion of the initial income level and the variables in the Z^i vector namely investment, schooling, trade openness, political instability and natural resources. Although the variables for schooling and political instability are statistically insignificant, all coefficients display their expected sign and are hence economically significant. Most notably, the variables that measure trade openness and natural resource abundance are significant at 1%.

Regression (3) in *Table 1* includes the corruption variable to regression (2) to account for the direct effects of corruption on growth. Relative to the first regression, the coefficient for corruption has dropped in significance. This relative insignificance of corruption on growth is, according to Pellegrini & Gerlagh (2004), due to the fact that a large part of the effect of corruption on growth is transmitted through other variables such as investment, openness, education and political stability, and here, natural resources. Their coefficients partly reflect the indirect effects for each transmission channel. However, unlike in Pellegrini & Gerlagh (2004), the non-zero coefficient on corruption in regression (3) suggests that the reduction in the corruption index leads to a substantial direct effect on the growth rate. This could imply the existence of other transmission channels which have not been identified.

In all the regressions, the coefficient capturing conditional convergence is statistically significant at 1%, consistent with the growth literature. The increase in (the absolute value of) the coefficient of $\ln(Y^i_{1970})$ in regressions (2) and (3) can be interpreted as an improvement of the identification of the steady state path of the economy. The large coefficient for trade openness is reminiscent of Sachs & Warner's (1995) findings, but quantitative results need to be interpreted with caution as there could be substantial bias in the estimates in *Table 1*, especially since corruption is found to be endogenous in the literature. Furthermore, the adjusted R^2 is low throughout all regressions.

V.2 Two-Stage Least Squares

In 2SLS estimation procedure with all transmission channels included, some or all of the coefficients take the wrong expected sign and lose significance due to large standard errors[11]. For brevity's sake, none of the 2SLS estimates are reported but a summary of the necessary 2SLS tests is presented in *Table 2* below.

2SLS	Instrument	n	First-stage t-stat	Second-stage Wu-Hausman F-Test	Second-stage Sargan's NR^2 Test
(1)	*legor_uk*	83	-0.63 (0.53)	0.75 (0.39)	-
(2)	*log_pop_dens_1500* *log_sett_mort*	49	1.52 (0.14) 1.94 (0.06)	4.06 (0.05)	6.51 (0.01)
(3)	*mal_46a*	83	2.87 (0.01)	0.00 (0.99)	-
(4)	*mal_46p*	83	2.95 (0.00)	0.64 (0.43)	-
(5)	*ethnic* *language*	80	2.31 (0.02) -1.59 (0.12)	4.89 (0.03)	5.26 (0.02)
(6)	*elf61*	79	1.12 (0.27)	5.11 (0.03)	-
(7)	*elf85*	80	0.82 (0.42)	5.54 (0.02)	-

Table 2 Summary of 2SLS tests on instruments[12]
P-values are in parenthesis

Corruption is only endogenous in the 2SLS regressions (2), (5), (6) and (7), but the instruments are weakly correlated, if at all, in the first-stage. Of all the instruments investigated only malaria rates are strongly correlated with corruption, but the second-stage fails the endogeneity test. The use of legal origins as an instrument fails both requirements at the first- and second-stage, which casts doubt on the robustness of the estimates in Pellegrini & Gerlagh (2004). Since all investigated instruments are weakly correlated with

[11] The same conclusion is reached when the natural resources variable is excluded.
[12] These tests were performed computationally by STATA v. 9.2 using the commands *ivendog* for the Wu-Hausman test and *overid* for Sargan's test for overidentifying restrictions.

corruption, the 2SLS estimates cannot be used for a quantitative assessment of the transmission channels. These findings are consistent with Shaw et al (2006) and add to the list of instruments that cannot be used.

V.3 Direct and Indirect Effects of Corruption

Due to the endogenous nature of corruption in the growth regression, OLS estimates will be biased. However, the OLS estimates are better than 2SLS estimates when only weak instruments are available. Since the coefficients are biased, quantitative interpretation is of little value. Nonetheless, the dependence of the variables in the Z^i vector on corruption based on equation (2) using OLS estimates[13] is presented in *Table 3.1* below.

Dependent Variable:	(4) *invest7000*	(5) *sec70*	(6) *openness7092*	(7) *pinstab7085*	(8) *natres70*
lgdp70	0.26	3.83***	0.03	-0.01	0.07***
	(0.69)	(1.38)	(0.04)	(0.02)	(0.02)
corrupt	**-2.02*** **	**-2.93*** **	**-0.14*** **	**0.01* **	**0.03*** **
	(0.33)	**(0.58)**	**(0.02)**	**(0.01)**	**(0.01)**
constant	24.26	-1.79	0.94	0.12	-0.57
N	129	98	102	114	105
Adjusted R^2	0.38	0.56	0.62	0.09	0.10

Table 3.1 Indirect transmission mechanisms as in equation (2)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

The coefficient estimates are larger in size and significance compared with Pellegrini & Gerlagh (2004) which is unsurprising, as there are potential endogeneity effects unaccounted for. The direct and indirect effects as in equation (3) are presented in *Table 3.2* below.

[13] In their main analysis, Pellegrini & Gerlagh (2004) presented OLS estimates and found it to be robust after adding other independent variables such as a democracy index, an OECD dummy and regional dummies.

Dependent Variable: *growth7000*	OLS (9)
lgdp70	-1.18*** (0.32)
$\hat{\mu}_1$ *(invest7000)*	0.05** (0.03)
$\hat{\mu}_2$ *(sec70)*	0.01 (0.02)
$\hat{\mu}_3$ *(openness7092)*	2.29*** (0.59)
$\hat{\mu}_4$ *(pinstab7085)*	0.03 (1.34)
$\hat{\mu}_5$ *(natres70)*	-3.96** (1.91)
corrupt	**-0.71*** (0.12)**
Constant	14.9
n	83
Adjusted R^2	0.42

Table 3.2 Growth Regressions as in equation (3)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

A comparison between the results of regression (9) and (3) shows that after taking into account of the indirect transmission mechanisms, the coefficient for corruption has become highly significant and has increased by more than a factor of six. The coefficient for corruption is almost twice as large as in Pellegrini & Gerlagh (2004)[14] which prove that their quantitative analysis is not robust. The results also differ significantly when the variable for natural resource abundance is excluded which means that it is not robust to the mere inclusion of more countries and an extended time-frame, despite similar variables and techniques employed[15]. Nonetheless, it still appears the case that corruption is the most important variable explaining growth when the

[14] -0.71 cf. -0.38
[15] See Appendix D.

transmission channels are taken into account. The size of the coefficient on the parameter for conditional convergence suggests that, holding the explanatory variables constant, an economy approaches its steady state at approximately 1.2% per year.

Table 3.3 below summarises the relative contribution of corruption and its transmission channels to the total effect on economic growth.

Transmission Channel	α_3 (Table 1)	β_2 (Table 2)	contribution to $\alpha_2 + \alpha_3\beta_2$	Relative contribution
corrupt	-	-	-0.13	17.6%
invest7000	0.05	-2.02	-0.11	15.3%
sec70	0.01	-2.93	-0.04	6.1%
openness7092	2.29	-0.13	-0.31	44.2%
pinstab7085	0.03	0.01	0.00	0%
natres70	-3.96	0.03	-0.12	16.8%
Total			-0.71	100%

Table 3.3 The relative importance of each Transmission Mechanisms as in equation (3)

Interestingly, the direct impact of corruption on growth is similar in magnitude with Pellegrini & Gerlagh (2004)[16]. However, evidence from *Table 5.4* in Appendix D implies the finding is spurious. The main effects of corruption are transmitted through trade, natural resources, investment and education in the order of importance; and it has a negligible effect on political instability. The effect of corruption on trade openness is 44.2%, which is almost three times as large as investment, which is in stark contrast with previous empirical literature that emphasised the importance of corruption on investment[17]. This is further reinforced in *Table 5.4* in Appendix D. Given the relative lack of theoretical and empirical evidence in previous literature on the importance of trade openness on corruption, this supports the skeptical view

[16] See Pellegrini & Gerlagh (2004): 17.6% cf. 19%

[17] Mauro (1995)

of Rodriguez and Rodrik (1999) that the available measures of trade openness tend to overstate the logical evidence in its favour. Furthermore, there is evidence that the discovery of natural resources will exert a negative impact on growth in the presence of corrupt institutions in a country. This supports recent work by Leite & Weidmann (1999) and Sala-i-Martin & Subramanian (2003) who attribute natural resource curse to corruption.

V.4 Long-Term Effect of Corruption on the Transmission Channels

	invest7000	*sec70*	*openness7092*	*pinstab7085*	*natres70*
Direct effect	-2.02	-2.93	-0.14	0.01	0.03
Indirect income effect	-0.43	-2.32	-0.02	0.01	-0.04
Long-term effect	**-2.45**	**-5.25**	**-0.16**	**0.02**	**-0.01**

Table 4.1 The long term effect of corruption on the transmission variables as calculated in equation (4)

	(10) *invest7000*	(11) *sec70*	(12) *openness7092*	(13) *pinstab7085*	(14) *natres70*
corrupt	**-2.11*****	**-4.13*****	**-0.15*****	**0.02*****	**0.01**
	(0.24)	**(0.39)**	**(0.01)**	**(0.01)**	**(0.01)**
Constant	26.83	36.21	1.21	-0.01	0.10
N	129	98	102	114	105
Adjusted R^2	0.38	0.52	0.63	0.10	0.01

Table 4.2 The indirect transmission channels calculated as in equation (5)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

The long-term effects of corruption on the transmission channels, investment, schooling, trade openness, political instability and natural resources are given in *Table 4.1* and *Table 4.2* above.

The estimates of the coefficients in equation (4) are slightly larger, in absolute value, than those from equation (5) as in Pellegrini & Gerlagh (2004), due to the fact that equation (4) implicitly assumes the level of corruption to be constant over time, and thus gives too much weight to the future effect of present corruption levels. At the same time, the small gap between the two sets of values strongly supports the relative persistence over time of corruption levels. Finally, although the results imply a significant long-term effect of corruption on those variables that are drivers of economic growth (investment, trade openness, schooling and political stability), the evidence also suggests that corruption does not have a long-term effect on natural resources on balance; natural resources can have a positive or negative effect on economic growth.

Section VI
Limitations & Extensions

The empirical analysis of corruption is difficult due to problems of measurement as well as endogeneity bias. Corruption can have different interpretations in different cultures, and what is usually considered a gift in one, might be interpreted as a bribe in another. The perceptions' of corruption acts only as a proxy to actual corruption, but can be biased even further if perceptions are backward-looking. Cross-sectional studies take corruption to be constant, and hence ignore variations in corruption over time. The absence of good instruments seems to suggest a cul-de-sac for 2SLS methods to correct for reverse causality in the corruption literature. Two avenues for future research are suggested.

First, endogeneity bias can also be dealt with by moving to a panel from pure cross-sectional data. The persistence of corruption suggests that a five- or ten-year balanced panel is preferable. Also, using panel data analysis makes it easier to identify how structural changes affect corruption over time, and to assess whether or not poverty breeds corruption and vice versa. The changes in sample and methodology mean that Transparency International data cannot be used (yet) for comparisons over time, but this might be possible with other sources of data such as the International Country Risk Guide published by Political Risk Services which must be purchased.

Second, since the results of large-scale econometric studies like this tend to disguise the complexity of corruption, evidence from micro-level studies is necessary to suggest country-specific ways of combating corruption. Transparency International noted that Finland, Iceland and New Zealand shared the top score in the latest index, providing a good precedent that corruption can be almost eradicated. What are the economic, political and cultural conditions in place? Even poorer countries such as Chile, Barbados, Uruguay, Jordan and Botswana score well in the index, which implies that the downward spiral of poverty and corruption can be overcome.

Section VII
Conclusions

Overall, while two-stage least-squares estimation would be useful to account for the endogeneity of corruption, all of the instruments are found to be weak and quantitative results are shown to be unreliable since the effects of reverse causality are insufficiently accounted for. The persistent interest in the corruption and growth literature after Mauro (1995) highlights the many shortcomings of his initial analysis and suggests that the existing approaches still have not got it "quite right".

The danger of cross-sectional analysis is the increased possibility of endogeneity bias. In a certain cross-section, if high levels of corruption are also associated with low levels of GDP, this is often interpreted as causation (i.e. high corruption will cause low growth or low growth will cause high corruption). It is also often interpreted longitudinally (i.e. a reduction in corruption levels will lead to an increase in GDP). Neither deduction is strictly correct based on a cross-sectional analysis. The only correct inference in this situation is that high GDP countries have low corruption levels and the issue of reverse causality remains pressing.

(4713 words)

References

Acemoglu, D., Johnson, S. & Robinson, J.A., 2001, "The Colonial Origins of Comparative Development: An Empirical Investigation", *American Economic Review*, Vol. 91, pp. 1369-1401.

Acemoglu, D., Johnson, S. & Robinson, J.A., 2002, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution", *Quarterly Journal of Economics*, Vol. 117, pp. 1231-1294.

Acemoglu, D. & Johnson, S., 2005, "Unbundling Institutions", *Journal of Political Economy*, Vol. 113, 949-995

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003, "Fractionalization", *Journal of Economic Growth*, Vol. 8 pp. 155-194

Gallup, J. L., Sachs, J.D., Mellinger, A., 1999, "Geography and Economic Development", CID Working Paper No. 1, March 1999 <http://www.cid.harvard.edu/ciddata/ciddata.html>

Heston, A., Summers R. & Aten, B., 2006, "Penn World Table Version 6.2", *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.*

La Porta, R., Lopez-de-Silanes, F., Shliefer, A. & Vishny, R., 1999, "The Quality of Government," *Journal of Law, Economics and Organization*, Vol. 15(1) pp. 222-279

Leite, C. and Weidmann, J., 1999 "Does Mother Nature Corrupt? Natural Resources, Corruption and Economic Growth", IMF *Working Paper No. 85*, Washington, DC: International Monetary Fund.

Levine, R., & Renelt, D., 1992, "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review* Vol. 82(4) pp. 942-963

Mauro, P., 1995, "Corruption and Growth", *Quarterly Journal of Economics* Vol. 110(3) pp. 681-712

Mauro, P., 1998, "Corruption and the Composition of Government Expenditure", *Journal of Public Economics* Vol. 69 pp.263-279

Pellegrini, L. & Gerlagh, R., 2004, "Corruption's Effect on Economic Growth and Its Transmission Channels" *Kyklos*, Vol. 57(3) pp.429-456

Rodriguez, F. and Rodrik, D., 1999, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." *CEPR Discussion Papers Series*, No.2413

Roeder, P.G., 2001, "Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985" February 16. <http//weber.ucsd.edu\~proeder\elf.htm>

Rose-Ackerman, S, 1999, *"Corruption and Government: Causes, Consequences and Reform"*, Cambridge University Press.

Sachs, J.D. & Warner, A.M., 1995, "Natural Resource Abundance and Economic Growth," *NBER Working Papers 5398*, National Bureau of Economic Research, Inc.

Sachs, J.D. & Warner, A.M., 1997, "Fundamental Sources of Long-Run Growth," A*merican Economic Review*, Vol. 87(2) pp.184-188.

Sachs, J.D. & Warner, A.M., 2001, "The Curse of Natural Resources," European Economic Review Vol. 45 pp.827-838.

Sala-i-Martin, X., 1997, "I Just Ran Two Million Regressions", American Economic Review, Vol. 87 (2) pp.178-183

Sala-i-Martin, X. & Subramanian, A., 2003, "Addressing the Natural Resource Curse: An Illustration from Nigeria," *NBER Working Papers 9804*, National Bureau of Economic Research, Inc.

Shaw, P., Katsaiti, M.-S. & Jurgilas, M., 2006, "Corruption and Growth under Weak Identification", *University of Connecticut Department of Economics Working Paper Series (17)*

Shleifer, A. & Vishny, R., 1993, "Corruption" *Quarterly Journal of Economics,* Vol. 108(3) pp.599-617

Treisman, D., 2000, "The causes of corruption: A cross-national study", *Journal of Public Economics*, Vol. 76 pp. 399-457

Appendix A:

Countries in the Sample

Algeria	Greece	Niger
Argentina	Grenada	Nigeria
Australia	Guatemala	Norway
Austria	Guinea	Oman
Bahrain	Haiti	Pakistan
Barbados	Honduras	Panama
Belgium	Hong Kong	Papua New Guinea
Belize	Hungary	Paraguay
Benin	Iceland	Peru
Bhutan	India	Philippines
Bolivia	Indonesia	Poland
Botswana	Iran	Portugal
Brazil	Iraq	Qatar
Burkina Faso	Ireland	Romania
Burundi	Israel	Rwanda
Cambodia	Italy	Saudi Arabia
Cameroon	Jamaica	Senegal
Canada	Japan	Sierra Leone
Central African Republic	Jordan	Singapore
Chad	Kenya	South Africa
Chile	Kiribati	Spain
China	Korea, Republic of	Sri Lanka
Colombia	Kuwait	Sudan
Congo, Dem. Rep.	Laos	Suriname
Congo, Republic of	Lesotho	Swaziland
Costa Rica	Luxembourg	Sweden
Cote d`Ivoire	Macao	Switzerland
Cuba	Madagascar	Syria
Cyprus	Malawi	Taiwan
Denmark	Malaysia	Tanzania
Dominica	Mali	Thailand
Dominican Republic	Malta	Togo
Ecuador	Mauritania	Trinidad & Tobago
Egypt	Mauritius	Tunisia
El Salvador	Mexico	Turkey
Equatorial Guinea	Mongolia	Uganda
Ethiopia	Morocco	United Arab Emirates
Finland	Mozambique	United Kingdom
France	Namibia	United States
Gabon	Nepal	Uruguay
Gambia, The	Netherlands	Venezuela
Germany	New Zealand	Zambia
Ghana	Nicaragua	Zimbabwe

Total number of countries: 129

Appendix B:
Variable Definitions and Data Sources

The data were gathered mainly from the Penn World Tables 6.2 http://pwt.econ.upenn.edu/ (PWT) and the Harvard Centre for International Development http://www.cid.harvard.edu/ciddata/ (Harvard CID). Data for corruption is gathered from Transparency International http://www.transparency.org/

growth7000	Average annual growth rate of real per capita GDP over the period 1970-2000, PPP adjusted: $G^i = (1/T)\ln(Y_T^i / Y_0^i)$ where $t_0 = 1970$ & $t_T = 2000$. Source: PWT 6.2 using variable *rgdpch*
lgdp70	Natural logarithm of real per capita GDP in 1970, PPP adjusted. Source: PWT 6.2 using variable *rgdpch*
invest7000	Average share of gross domestic investment in GDP 1970-2000. Source: PWT 6.2 using variable *ki*
openness7092	Proportion of years a country has been open in 1970-1992 Source: Sachs & Warner (1995) from Harvard CID using variable *open*
sec70	Gross secondary school enrolment in 1970. Source: Barro & Lee (1993) from Harvard CID using variable *sec70*
pinstab7085	Average number of assassinations per million of inhabitants and revolutions between the period 1970 and1985. Source: Barro & Lee (1993) from Harvard CID using variables *pinstab*xx
natres70	Share of exports of natural resources in GNP in 1970, measured in %. Natural resource exports are the sum of exports of agricultural raw materials, foods, fuels, and ores and metals. Source: Sachs & Warner (1997) from Harvard CID using variable *sxp*
corrupt	The CPI represents a 'poll of polls' from individual surveys of businessmen or local populations of the relevant countries as well as ratings by economic risk analysts and country experts.

	The values of the 2006 CPI index range between 0 & 10 where a low (high) value indicates high (low) levels of corruption. This is subtracted from 10, so that an increase in the index will have the intuitive meaning of an increase in corruption. Source: Transparency International

Investigated Instruments

legor_uk	Dummy variable for British legal origin of each country. Source: La Porta et al. (1999)
log_pop_dens_1500	Natural log of total population divided by total arable land in 1500. Source: Acemoglu, Johnson & Robinson (2002).
log_sett_mort	Natural log of the mortality rate faced by European settlers at the time of colonization. Source: Acemoglu, Johnson & Robinson (2001)
mal46a	% of country area with malaria, 1946. Source Gallup, Sachs & Mellinger (1999)
mal46p	% of 1995 population living in areas with malaria, 1946.Source Gallup, Sachs & Mellinger (1999)
ethnic	The values of the index range between 0 & 1 where a low (high) value indicates low (high) rates of ethnic fractionalization. Source: Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg (2003)
language	The values of the index range between 0 & 1 where a low (high) value indicates low (high) rates of linguistic fractionalization. Source: Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg (2003)
elf61	Ethnolinguistic Fractionalization 1961. Source Roeder (2001)
elf85	Ethnolinguistic Fractionalization 1985. Source Roeder (2001)

Appendix C:

Summary Statistics

Variable	Number of observations	Mean	Standard Deviation	Minimum	Maximum
corrupt	129	5.66	2.281	0.4	8.2
lgdp70	129	8.11	1.087	5.426	11.109
growth7000	129	1.573	1.969	-4.551	6.935
invest7000	129	14.907	7.745	2.420	43.164
openness7092	102	0.382	0.434	0	1
sec70	98	13.850	13.693	0.1	63.9
pinstab7085	114	0.099	0.138	0	0.618
natresgnp70	105	0.152	0.153	0.006	0.886
Instruments					
legor_uk	129	0.326	0.470	0	1
log_pop~1500	80	0.434	1.580	-3.912	4.610
log_sett_mort	68	1.947	0.481	0.932	2.684
mal46a	126	0.651	0.407	0	1
mal46p	126	0.688	0.403	0	1
ethnic	128	0.452	0.263	0.002	0.930
language	123	0.409	0.294	0.002	0.923
elf61	119	0.471	0.272	0.006	0.909
elf85	122	0.479	0.277	0.007	0.984

Appendix D:
Further Econometric Results

This section presents the results that broadly replicate[18] the study by Pellegrini & Gerlagh (2004) with the use of more countries from the 2006 CPI index. They showed quantitative evidence to be robust after adding other independent variables such as a democracy index, an OECD dummy and regional dummies. However as shown here, the mere inclusion of more countries proves otherwise, thus supporting the conclusion in the main text. Quantitative results are unreliable due to the effect that endogeneity cannot be accounted for.

Dependent variable: *growth7000*	(15)	(16)	(17)
lgdp70	-0.89***	-0.74***	-0.87***
	(0.19)	(0.26)	(0.29)
invest7000		0.05*	0.05*
		(0.03)	(0.03)
sec70		0.01	0.01
		(0.02)	(0.02)
openness7092		2.55***	2.19***
		(0.49)	(0.59)
pinstab7085		-0.85	-0.51
		(1.24)	(1.28)
corrupt	**-0.64*****		**-0.16**
	(0.09)		**(0.15)**
constant	12.45	5.66	7.82
N	129	88	88
Adjusted R^2	0.27	0.36	0.36

Table 5.1 Growth regressions based on equation (1)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

[18] There are minor differences with regard to the time-frame employed and variables used. See Appendix B and Pellegrini & Gerlagh (2004)

Dependent Variable:	(18) *invest7000*	(19) *sec70*	(20) *openness7092*	(21) *pinstab7085*
lgdp70	0.26	3.83***	0.03	-0.01
	(0.69)	(1.38)	(0.04)	(0.02)
corrupt	**-2.02*****	**-2.93*****	**-0.14*****	**0.01***
	(0.33)	**(0.58)**	**(0.02)**	**(0.01)**
constant	24.26	-1.79	0.94	0.12
N	129	98	102	114
Adjusted R^2	0.38	0.56	0.62	0.09

Table 5.2 Indirect transmission mechanisms as in equation (2)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

growth7000	(22)
lgdp70	-0.76***
	(0.27)
$\hat{\mu}_1$ *(invest7000)*	0.05*
	(0.03)
$\hat{\mu}_2$ *(sec70)*	0.01
	(0.02)
$\hat{\mu}_3$ *(openness7092)*	2.19***
	(0.59)
$\hat{\mu}_4$ *(pinstab7085)*	-0.51
	(1.28)
corrupt	**-0.59*****
	(0.11)
Constant	10.9
N	88
Adjusted R^2	0.32

Table 5.3 Growth Regressions as in equation (3)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients

Transmission Channel	α_3 (Table 1)	β_2 (Table 2)	contribution to $\alpha_2 + \alpha_3\beta_2$	Relative contribution
Corrupt	-	-	-0.16	27.6%
Invest7000	0.05	-2.02	-0.10	16.3%
sec70	0.01	-2.93	-0.02	3.4%
openness7092	2.19	-0.14	-0.30	51.4%
pinstab7085	-0.51	0.01	0.01	1.3%
Total			-0.59	100%

Table 5.4 The relative importance of each Transmission Mechanisms as in equation (3)

	invest7000	*sec70*	*openness7092*	*pinstab7085*
Direct effect	-2.02	-2.93	-0.14	0.01
Indirect income effect	-0.19	-2.95	-0.02	0.01
Long-term effect	**-2.22**	**-5.88**	**-0.16**	**0.02**

Table 5.5 The long term effect of corruption on the transmission variables as calculated in equation (4)

	invest7000	*sec70*	*openness7092*	*pinstab7085*
corrupt	**-2.11***** **(0.24)**	**-4.13***** **(0.39)**	**-0.15***** **(0.01)**	**0.02***** **(0.01)**
Constant	26.83	36.21	1.21	-0.01
N	129	98	102	114
Adjusted R^2	0.38	0.52	0.63	0.10

Table 5.6 The indirect transmission channels calculated as in equation (5)
Superscripts *, **, *** correspond to 10%, 5%, 1% significance, respectively.
Standard errors are in parenthesis under the coefficients