Discussion Paper Series

Corruption and Development: A test for Non-linearities

By

M. Emranul Haque and Richard Kneller

Centre for Growth and Business Cycle Research, Economic Studies,
University of Manchester, Manchester, M13 9PL, UK

October 2007
Number 092

Corruption and Development:
A test for Non-linearities*

By

M. Emranul Haque [a, **] and Richard Kneller [b]

[a] Centre for Growth and Business Cycles Research, Economic Studies, The University of Manchester
[b] The Leverhulme Centre for Research in Globalisation and Economic Policy, School of Economics, University of Nottingham, UK

October 2007

Abstract

Using different combinations of culture, development and openness to international trade, we test the variability in the incidences of corruption at different stages of development or in other words the non-linearities in the relationship between corruption and development. We employ formal threshold model developed by Hansen (2000), and *unlike the existing literature*, we find that: (1) non-linear models that search for the break points in the relationship between corruption and development are statistically preferable than linear regressions; (2) the effect of development at any stage is much lower than that has been suggested by studies using linear regressions approach; (3) both culture and openness do not affect corruption directly; rather they have an effect on the location of break points in the relationship between corruption and development.

JEL Codes: D73, H26, O11

Keywords: Corruption, Development, Culture, Openness, Non-linearity

---

* We wish to thank the Leverhulme Trust (Programme Grant F114/BF) for financial support.

[b] Corresponding author: M Emranul Haque, Centre for Growth and Business Cycles Research, Economic Studies, The University of Manchester, UK, M13 9PL, Tel: +44 (0) 161 275 4829; Email: emranul.haque@manchester.ac.uk

1. Introduction

Over the past decade there has been a renewed interest among researchers, national governments and international organisations in the economic causes and consequences of corrupt behaviour within government bureaucracies. This has been motivated by a strengthening conviction that good quality governance is essential for sustained economic development and that corruption in the public sector is a major impediment to growth and prosperity. Indeed, increasingly development assistance is being made conditional that attempts at reducing corruption and establishing good governance are made. [1]

Recent innovations at the theoretical and empirical level have allowed this conviction to be tested formally, and there is now a large body of evidence to support the idea that corruption and development are strongly related. If one was to summarise the lessons that have been learnt from this economics literature the first, and perhaps the least surprising lesson, would be that there is a strong negative correlation between corruption and the level of development. This correlation is evident both anecdotally and at a more formal level. For example, outside of Italy, the often-mentioned outlier in this relationship, there are no examples of rich countries with high levels of corruption (corruption levels in Italy are in fact below the average for all countries), and no poor countries with low levels of corruption. Or more formally: if the level of corruption is regressed on the level of GDP per capita then, depending on the measure of corruption used, it is possible to explain between 50 to 73 per cent of the variation in the corruption index with this variable alone (Triesman, 2000). Evidence also exists for reverse causation between corruption and development; Mauro (1995) and Knack & Keefer (1995) have for example, found evidence of a significant relationship between higher levels of corruption and lower rates of both growth of output and capital investment.

However, while the negative correlation between corruption and the level of development is well established in empirical studies,[2] the assumed linearity falls short of explaining two additional stylised facts about the relationship between corruption and

[1] See for example the Extractive Industries Transparency Initiative set by the UK Government; the OECD 1997 Convention on Combating Bribery of Foreign Officials in International Business Transactions; or the recent addition (June 2004) of a 10^{th} principal of corporate behaviour to cover corruption by the UN.
[2] See, among others, Mauro (1995) and Knack & Keefer (1995), Ades & Di Tella (1999), La Porta et al. (1999), Triesman (2000), Persson et al. (2001).

development. *First*, corruption levels are highly persistent across time. Some countries remain trapped in a 'bad' equilibrium, or what we might label corruption clubs, characterised by pervasive corruption and low development, whereas others end up in a low corruption, high-income equilibrium (see Mauro, 2004; Blackburn *et al.*, 2006). A *second* is that there is much greater diversity in corruption levels among middle-income countries compared to the high and low corruption countries, an indication of multiple equilibria (Blackburn *et al.*, 2002, 2005).

In this paper we search for evidence of a non-linear relationship between corruption and development as well as its causes and its empirical consequences using cross-country data for 87 countries from 1980 to 2003. Using the predictions from theory, principally those of multiple equilibria, we search for break points in the relationship of development with corruption using the threshold methodology of Hansen (2000). These models search over the values of an identifier variable to determine where the relationship between two variables (in our case, corruption and development) changes. We then test for its significance and provide a confidence interval around this point. Three identifier variables are used in the paper, development itself, culture and openness to international trade. In all cases we find that whether alone, or in combination with each other, they are capable of finding evidence of non-linearities. That said, despite attempts, we cannot establish statistically which is the preferred non-linear model and therefore the most likely determinant of the non-linearities we find. Initial investigations suggest that this is because countries are placed in similar regimes (they lie above and below particular threshold points) by each of the identifying variables. We therefore interpret the results as consistent with the broad class of theoretical models with multiple-equilibria and strongly suggest their further theoretical and empirical investigation.

The search for non-linearity also generates some strong conclusions about existing empirical work imposing a linear relationship, even those controlling for endogeneity, and the policy conclusions that might be drawn from these models. *First*, the effect of changes of development on corruption is weaker in the non-linear model. Changes to development do not lead to as large a change in corruption as predicted by the simple model, no matter whether the existing level of development is high, low or intermediate. *Second*, we find strong statistical evidence that the inclusion of variables such as culture and openness, found to have a significant effect on corruption in the linear model, may be misspecified. Culture and openness instead, have an effect on the location of the break points in the relationship of development with corruption. That is, when these variables are used as the identifier variable, and are therefore *not* included on the right hand side of the regression equation, we can establish statistical preference over a linear regression model in which they are. This has striking implications for existing empirical work and the future direction of empirical investigation into the determinants of corruption.

The remainder of the paper is organised as follows. In Section 2 in order to motivate the empirical evidence contained in the paper we provide a review of the literature on multiple-equilibria. Section 3 then describes the data being used and provides some initial evidence in support of the idea of thresholds in the data, while Section 4 describes the empirical methodology. Section 5 presents the initial set of empirical results on the preference of non-linear models over linear one, while Section 6 tests for robustness using alternative identifiers. Finally in Section 7, we offer some concluding remarks.

2. Literature Review

Prior to the early 1990s, the lack of reliable data on corruption meant that little was known about the true effects (if any) of bureaucratic malfeasance on economic development. Given this, it was possible to entertain seriously the idea that corruption might actually be conducive to growth and prosperity, based on the argument that corruption (bribes) may help to circumvent cumbersome regulations (red tape) in the bureaucratic process (e.g., Huntington 1968; Leff 1964; Leys 1970; Lui 1985).

It is now generally accepted that efficiency-enhancing and growth-promoting corruption is very much the exception, rather than the rule. The contemporary wisdom is that corruption is typically bad for development due to its adverse effects on the incentives, prices and opportunities that private and public agents face. This consensus of opinion is based on a large body of empirical literature – pioneered by Mauro (1995) – that has flourished over recent years due to the publication of several cross-country data sets that are widely regarded as providing reliable measures of corrupt activity.[3] This literature,

[3] These data sets, or corruption indices, have been complied by various international organizations (most notably Business International Corporation, Political Risk Services Incorporated and Transparency

both formal and anecdotal, has yielded a number of important stylised facts, which we summarize briefly as follows.

First, there appears to be a robust (and significant) negative correlation between the level of corruption and economic growth. According to Mauro (1995), the principal mechanism through which corruption affects growth is a change in private investment: an improvement in the corruption index by one standard deviation is estimated to increase investment by as much as 3 percent of output. Others have found similar evidence of significant adverse effects of corruption on growth (e.g., Gyimah-Brempong 2000; Knack and Keefer 1995; Li et al. 2000).

Second, there is evidence that the relationship between corruption and growth is two-way causal: bureaucratic rent-seeking not only influences, but is also influenced by, the level of development. In a thorough and detailed study by Treisman (2000), rich countries are generally rated as having less corruption than poor countries, with as much as 50 to 73 percent of the variations in corruption indices being explained by variations in per capita income levels. These findings, supported in other studies (e.g., Ades and Di Tella 1999; Fisman and Gatti 2002; Montinola and Jackman 1999; Paldam 2002; Rauch and Evans 2000), suggest that cross-country differences in the incidence of corruption owe much to cross-country differences in the level of prosperity.

A third stylised fact that could be drawn from the existing literature would be that corruption levels are highly persistent across time. Some countries remain trapped in a 'bad' equilibrium, or what we might label corruption club, characterized by pervasive corruption and low development, whereas others end up in a low corruption, high-income equilibrium (e.g., Bardhan 1997; Sah 1991; Mauro, 2004, Blackburn *et al.*, 2006). A cursory inspection of the data reveals that many of the most poor and corrupt countries in the past are among the most poor and corrupt countries today (Blackburn *et. al.*, 2006).[4] This

International) using survey questionnaires sent to networks of correspondents around the world. These corruption indices rank countries according to the extent to which corruption in public office is perceived to exist. While differing in their precise construction, the indices are very closely correlated with each other, lending support to the contention that they provide reliable estimates of the actual extent of corruption activity.

[4] Examples include Bangladesh, Cameroon, India, Indonesia, Kenya, Nigeria, Pakistan and Uganda. According to the data from Transparency International, these belong to a set of countries that have displayed little, or no, improvement in their corruption and growth records since the early 1980s.

conjures up the idea of poverty traps and the notion that some countries may be drawn into a vicious circle of low growth and high corruption, from which there is no easy escape.

A final feature of the data – that has received attention quite recently (e.g., Blackburn *et. al.*, 2002, 2005) – is the much greater diversity in the incidence of corruption among middle-income countries, compared to the uniformly high levels of corruption among low-income countries and the uniformly low levels of corruption among high-income countries.

At the micro level, several authors have pointed to strategic complementarities as a major factor in determining this feature of the data. The individual incentive to be corrupt depends on the behaviour of other agents in the economy and it is this corruption culture that gives rise to multiple equilibria. If corruption is pervasive then agents will choose to be corrupt also, whereas if others are generally honest then agents will choose not to be corrupt. A number of similar mechanisms have been used to generate this result. Andvig and Moene (1990) for example, assume that the probability a corrupt official will be reported to higher authorities is a decreasing function of the proportion of his colleagues who are also corrupt, while Sah (1991) explains differences in crime participation rates across otherwise similar societal groups on the basis of a learning model in which it is easier to observe members of one's own group, and Tirole (1996) the interaction between the reputation of a group and its individual members. Similarly Murphy et al (1993) have shown that aggregate returns to rent-seeking relative to market production is increasing until the proportion of rent-seekers to producers reaches an upper limit, while Putman (1993) has argued that a 'tragedy of commons' may explain the institutional and economic failure of some Italian regions. [5]

At the macro-theoretical level, various models have been developed to account for this correlation. In one strand, the approach has been to treat corruption as exogenous and to focus on its role in influencing growth by diverting the resources away from productive to unproductive sectors (Ehrlic and Lui, 1999), by reducing the fraction of savings channelled to investment (Sarte, 2000), or by lowering the amount and quality of public

[5] See also Cadot (1987), Dawid and Feichtinger (1996), and Huang and Wu (1994).

infrastructure and services supplied to the private sector (Del Monte and Papagni, 2001). In another strand, the approach has been to model the incentive to engage in corrupt activity as an endogenous outcome of the growth process itself and to focus on the co-evolution of real and corrupt activity (e.g., Blackburn *et al* 2002, 2005, 2006; Mauro, 2004) – utilizing the phenomenon of either 'taxation and bribery', or, 'resource misallocation and embezzlement'. These models can produce multiple equilibria and explain why some countries remain trapped with high corruption – low-income equilibrium, and why there is diversity in the incidence of corruption among the middle-income countries.

While theoretical models have been able to explain all stylized facts mentioned above, formal empirical investigations into the relationship between corruption and development have concentrated their efforts on the first two stylised facts. Corruption and development have been regressed alternately on one another as researchers have sought to examine separately the key determinants of each. In both cases the typical approach has been to supplement simple ordinary least squares estimation with instrumental variables estimation as a means of correcting for any potential endogeneity. The results obtained indicate clearly that corruption and development are influenced strongly by each other.

While consistent with simple theories, the recent models of Blackburn *et. al.*, (2002, 2005, 2006) suggest that there is more to the issue of simultaneity than simply the interdependence of variables. Simultaneity might also lead to non-linearity as a result of endogenous thresholds and multiple development regimes.

In this paper, building on the observation that the direction of causation between corruption and development cannot be separated from non-linearity, we search for empirical evidence of these non-linearities between corruption and development in cross-country data. Two questions naturally pose themselves at this point. *First*, what is the form these non-linearities might take? Should they be modelled using higher order terms of the independent variables or as break points in a linear relationship for example. *Second*, can we establish statistical preference for one form of non-linearity over another? We make some progress on both fronts in this paper providing a strong preference for non-linear over linear models, although the results are suggestive that further empirical investigation is required.

On the first question: consistent with the recent theoretical models we model the non-linearities as break points in the estimated relationship between corruption and development, using the endogenous search techniques outlined by Hansen (2000) to determine the critical values for any thresholds and the confidence interval surrounding these thresholds. This same methodology has been usefully adopted in a number of other applied contexts that include studies of development and growth (e.g., Chemlarova and Papageorgiou 2005; Girma et al. 2003; Masanjala and Papageorgiou 2004; Papageorgiou 2002). This methodology is of a similar spirit to alternative approaches such as regression trees (see for example Durlauf et al., 2001 and Johnson & Takeyama, 2001). We detail the empirical methodology more fully in Section 4.

3. Data Sample and Characteristics

To measure corruption we use the index published by Transparency International for the period 1980 to 2003. This index, derived from survey questionnaires sent to networks of correspondents around the world, ranks countries according to the extent to which corruption in public (and political) office is perceived to exist. This index has a strong correlation with other sources on corruption, such as those by Business International Corporation and Political Risk Services Incorporated. For more detailed discussions and a comparison between measures see Ades and Di Tella (1997), Jain (1998), Tanzi and Davoodi (1997) and Treisman (2000).

Table 1: Summary Statistics

Time Period	*Obs.*	*Mean*	*Standard deviation*	*Minimu m*	*Maximu m*
Corruption	325	4.98	2.52	0.07	9.8
Ln(GDPpc)$_{t-1}$	325	8.23	1.56	4.97	10.73
Culture	325	4.25	1.16	0.58	6
Openness	282	70.22	53.24	10.05	384.99

In order to maximise the number of observations and to control for any business cycle effects we arrange the data into five non-overlapping time periods. These are 1980-1985, 1988-1992, 1995-1997, 1998-2000, and 2001-2003. There are missing years in this sequence as a result of the unavailability of the corruption index for all years.

Development is measured by GDP per capita in constant US$ and is taken from the WDI CDROM 2002. Table 1 reports some summary statistics of corruption and GDP.

The mean value of corruption across all countries and time periods is just under 5. The least corrupt country is Denmark with a corruption score of 0.07 (1998-2000) and the most corrupt country is Indonesia with a score of 9.8 (timper 1). There would appear to be two broad groups of countries however, those with a corruption score below three and those with medium to high corruption (a corruption score between 5 and 8).

Figure 1: Corruption Index and log GDP (2001-2003)

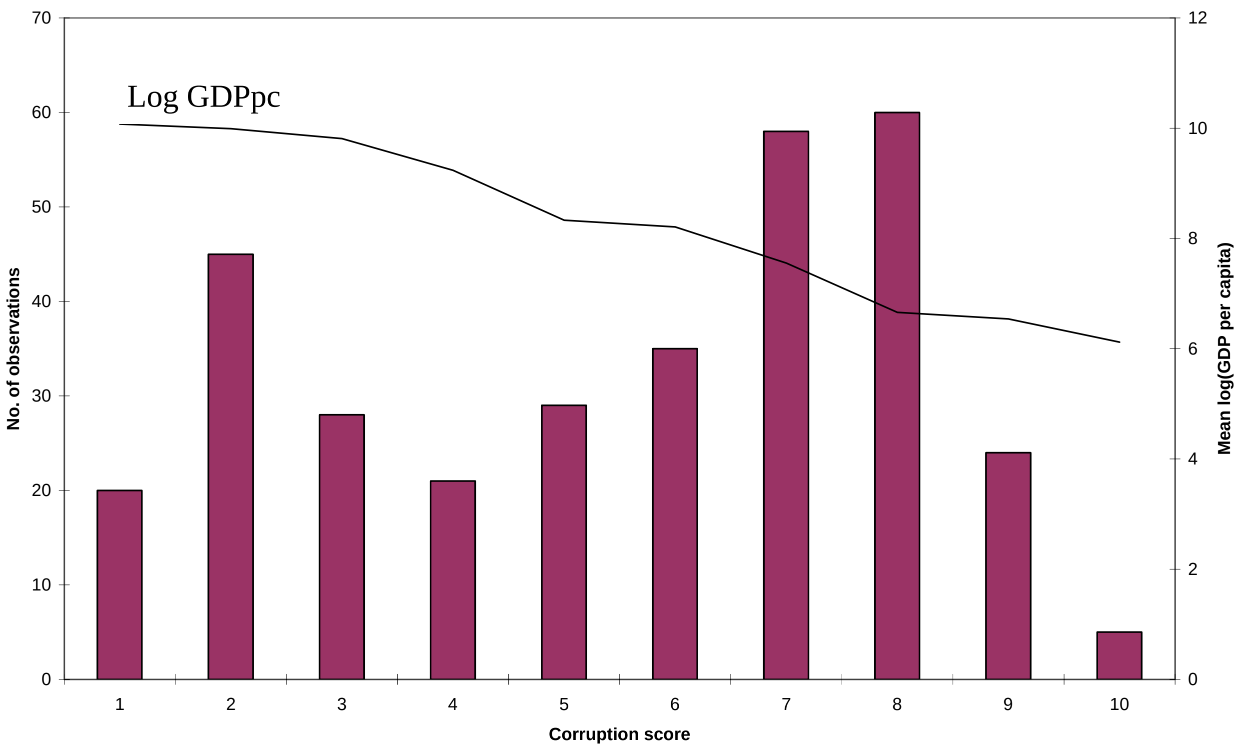
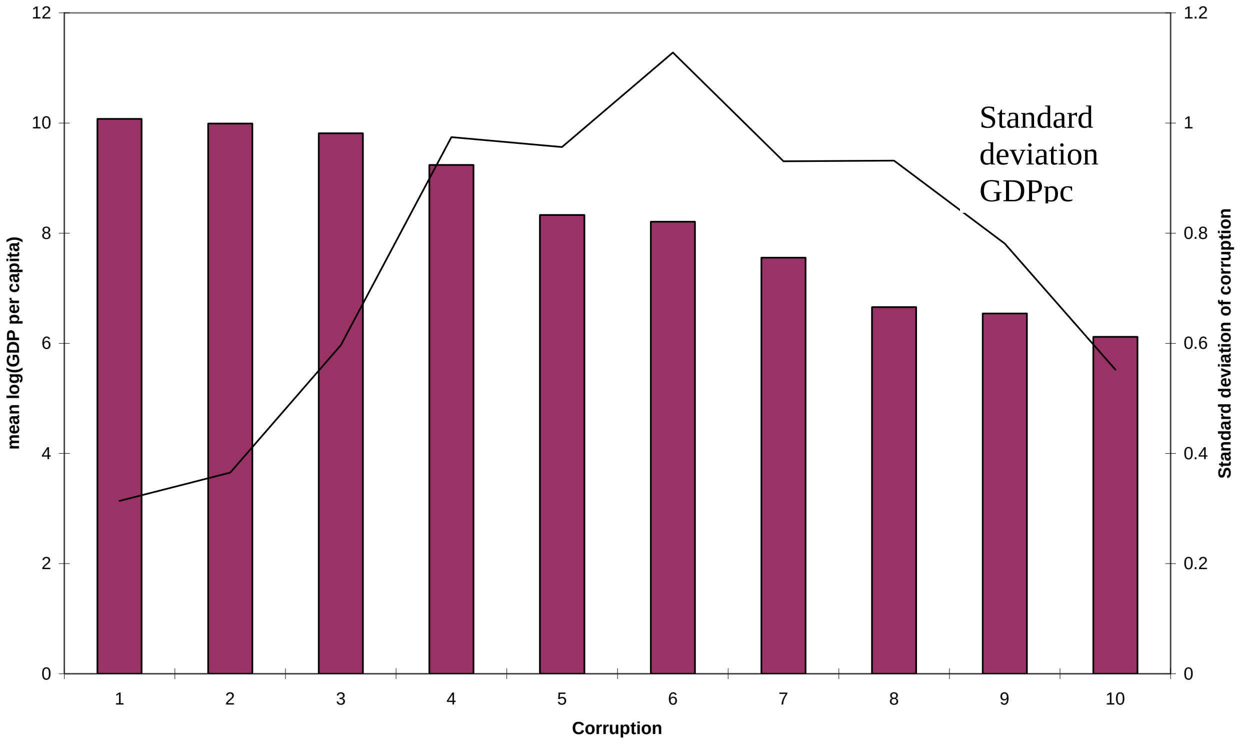


Figure 1 displays the information on the two variables of interest in the paper, corruption and the log of GDP per capita. The level of GDP per capita included in the chart corresponds to the average level of GDP for the observations included within each of the bars, while the bars themselves count the number of countries within the period 2001 to 2003 that fall within a given range of corruption. Two points are worth noting from this chart: firstly, the above-mentioned collection of countries around two levels of corruption, a low corruption and a medium/high corruption group. Secondly, that there is a clear negative trend between average GDP and corruption, providing a strong indication of the type of countries included in each column. Indeed over the sample the correlation between these two variables is -0.831. Countries with low corruption are typically those from OECD countries, whereas there is a greater mix amongst those with high corruption. For example, the countries with a corruption score between 7 and 8 include countries from Latin America, Eastern & Central Europe, Asia and Africa. This negative correlation is indicative of similar results using more formal regression analysis in the empirical literature.

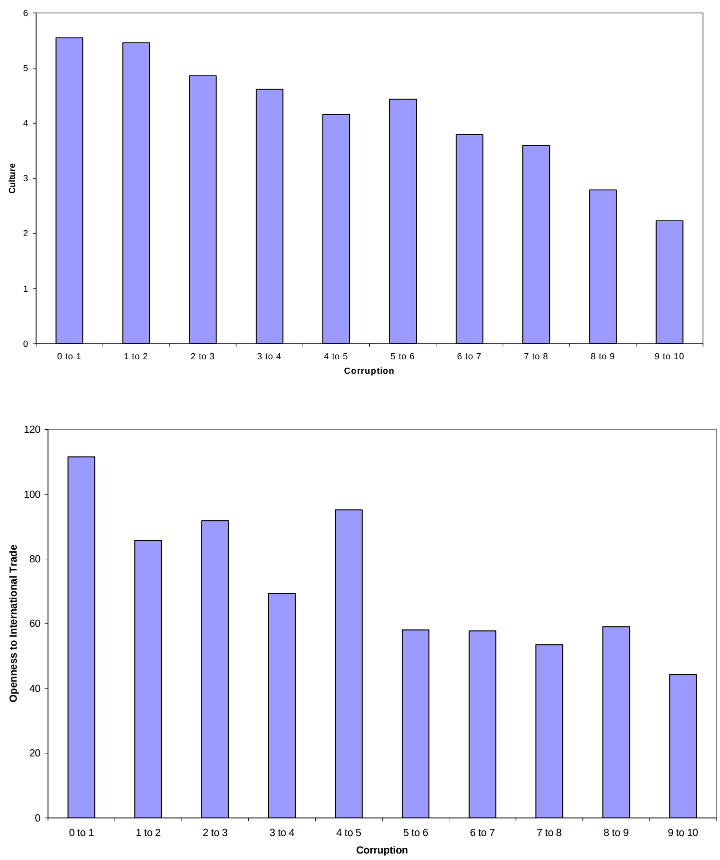
Figure 2: Corruption, GDP per capita and its standard deviation



Whilst corruption and GDP are clearly negatively correlated in Figure 1, unobserved is an increase in the standard deviation of GDP as we move from low to high levels of corruption. In Figure 2 we capture this hidden effect by plotting the average GDP per capita from Figure 1 against the standard deviation of GDP for each of the groupings of the corruption index (0-1, 1-2, 2-3 etc.). The x-axis in the Figure shows the value of the corruption index, the bars the average GDP per capita of countries within that range of corruption and the line the corresponding standard deviation of GDP per capita for countries with that level of corruption. The negative trend between corruption and GDP per capita remains evident in Figure 2, but now it is also clear that there is also a hump shape in the variation of GDP per capita amongst countries with a mid-range corruption score. The standard deviation of GDP per capita peaks at corruption levels between 5 and 6. This might be used as a first evidence of a possible threshold in the relationship.

Of the other variable used we measure culture using the sum of the indices for law and order, ethnic tensions and democratic accountability from International Country Risk Guide. We test the robustness of this measure to the addition of measures of military in politics and religion in politics from the same source. All of these indices are measured between 0 (low quality) and 6 (high quality). Trade exposure is measured as the ratio of exports and imports to GDP. Summary statistics for these variables are given in Table 1 above. The simple correlation of culture and openness with corruption is lower than that for development at –0.697 and –0.284 respectively, noticeably so in the case of openness. For completeness we plot the average score of the culture and openness variables at different levels of corruption in the same way as for Figures 1 and 2 above. In both cases a clear downward slope is evident.

Figure 3: Culture and Openness with Corruption



4. Formal Threshold Model

As mentioned in the literature review, there are theoretical models (Blackburn *et al.*, 2002, 2005) that predict that the correlation between corruption and development may vary with the level of development due to the influence of some other factors such as culture. Specifically, there are break points around which the relationship between corruption and development changes. We assume that the location of these break points, or thresholds, can be identified using information from the slope parameters in a regression of corruption and development. The testing procedure used to locate the position and significance of any thresholds in the GDP per capita/corruption relationship is the same as that used by Girma *et al* (2003) and Papageorgiou (2002). To understand the problems associated with the application of the testing procedure assume for simplicity that the corruption-GDP per capita relationship is captured by the single threshold equation given below:

$$CORR_{it} = \gamma X_{it} + \beta_1 \log(GDP)_{it} I(X_{it} \leq \alpha) + \beta_2 \log(GDP)_{it} I(X_{it} > \alpha) + \varepsilon_{it} \qquad (1)$$

where CORR is corruption, I(.) is an indicator function and β and γ are parameters to be estimated, where the indicator function and therefore the variable X is not directly included in the regression but instead helps to identify changes in the value of β_1. Evidence of a threshold in the effect of GDP per capita on corruption would be associated with a difference in the effect of GDP per capita on corruption above and below the critical value of the variable X, $\beta_1 \neq \beta_2$, where α identifies the break-point value of X. In estimation there are three steps: firstly, jointly estimate the threshold value α and the slope coefficients γ, β_1, and β_2. Secondly, test the null of no threshold (i.e. $H_0 : \beta_1 = \beta_2$) against the alternative hypothesis of a threshold model (i.e. $\beta_1 \neq \beta_2$) and thirdly, construct confidence intervals for α. We discuss what might be used for X in the indicator function below.

To estimate the parameters of the equation we use the algorithm provided in Hansen (2000), which searches over values of α sequentially until the sample splitting value $\hat{\alpha}$ is found[6]. Once found, estimates of γ, β_1 and β_2 are readily provided. The problem that arises in testing the null hypothesis of no threshold effect (i.e. a linear formulation) against the alternative of a threshold effect is that, under the null hypothesis, the threshold variable is not identified. Consequently, classical tests such as the Lagrange Multiplier (LM) test do not have standard distributions and so critical values cannot be read off standard distribution tables. To deal with this problem, Hansen (2000) recommends a bootstrap procedure to obtain approximate critical values of the test statistics, which allows one to perform the hypothesis test. We follow Hansen (2000) and bootstrap the p-value based on a likelihood ratio (LR) test.

If a threshold effect is found (i.e. $\beta_1 \neq \beta_2$), then a confidence interval for the critical level of X is generated. In this case one needs to test for the particular threshold value as: $H_o : \alpha = \alpha_0$. We require this confidence interval to be reasonably small, given countries within it cannot be confidently placed in either regime.

It should be noted that the test of the null hypothesis for forming the confidence interval is not the same as that for the second problem i.e. the test of no threshold effect. Under normality, the likelihood ratio test statistic $LR_n(\alpha) = n\frac{S_n(\alpha) - S_n(\hat{\alpha})}{S_n(\hat{\alpha})}$ is commonly used to test for particular parametric values. However, Hansen (2000) proves that when the endogenous sample-splitting procedure is employed, $LR_n(\alpha)$ does not have a standard χ^2 distribution. Consequently, he then derives the correct distribution function and provides a table of the appropriate asymptotic critical values and search over the remainder of the sample.[7]

Having identified the location of the first threshold the process is then repeated to find further thresholds. To do this we follow Papageorgiou (2002) and split the sample at the point of the first threshold value.

Blackburn *et al.*, (2002, 2005, and 2006) and Mauro (2004) suggest that equlibria between corruption and development are identified out of the interaction between culture and development. Our initial X variable is the log of culture interacted with development. We use a number of different variables as the identifying variable X however. This in turn raises the question of whether we can prefer any particular choice of identifying variable over another or indeed to the linear model. To answer this point we use a J-test (Davidson

---

[6] This is the value of α that minimises the concentrated sum of squared errors based on a conditional OLS regression.

and MacKinnon, 1981), where preference between these non-nested hypotheses is established on the basis of whether the maintained model can explain the variation of the data of the competing model (Greene, 2003).

5. Thresholds in Corruption

To provide a comparison with the results found allowing for non-linearities between corruption and development we report as model 1 the results from a linear regression with corruption regressed against the lag of the log of GDP per capita.[8] As expected we find a strong negative correlation between these two variables, indeed using just this variable we are able to explain over 70 per cent of the variation in corruption in the sample. According to the results an increase in the log of GDP by one standard deviation (1.56) is associated with a reduction in the level of corruption equal to 2.1 points on the corruption index. The standard deviation of corruption in the sample is 2.56.

Using the Hansen methodology we find evidence of three significant thresholds between corruption and GDP per capita, and all have well defined confidence intervals. The first lies at the 76th percentile of the GDPpc*culture distribution (p-value = 0.00[9], confidence interval 72nd – 78th percentiles[10]); the second at the 62nd percentile (p-value = 0.01, confidence interval 60th – 65th percentiles); and the third at the 13th percentile (p-value = 0.00, confidence interval 11th – 17th percentiles). In order to compare the effect of including thresholds we re-estimate regression equation 1 allowing the effect of GDP per capita on corruption to change within each of the ranges of the GDPpc*culture distribution identified above. This regression is reported as model 2 in Table 2, where I denotes the identifier variable listed in the second row of the table (in this case GDPpc*culture).

---

[7] See Table I on page 582 of Hansen (2000).

[8] Given our interest in the location of thresholds, we do not explore the issue of endogeneity in our regressions. However, in order to maximise the data points available to us (and to be consistent with the exiting empirical literature) we use the lag of GDP per capita rather than its contemporaneous value, an approach which is often used to minimise the effect of endogeneity on the results.

[9] All of the bootstrapped p-values in our endogenous threshold analysis are generated using 1000 bootstrap replications.

[10] Denoting the percentiles of the log of GDP by α, the 95% confidence interval for the threshold estimates is obtained by plotting the likelihood ratio sequence in α, LR (α), against α and drawing a flat line at the critical value (e.g. the 95% critical value is 7.35). The segment of the curve that lies below the flat line will be the confidence interval of the threshold estimate.

Several points are worth noting about this regression. Firstly, the fit of the regression improves compared to regression 1 from allowing the coefficient on GDPpc to vary. Second, in absolute value the relationship between GDP per capita and development is stronger in the regions above the bottom threshold, *i.e.* above the 13^{th} percentile of the identifying GDPpc*culture distribution, and strongest at the highest combinations of development and culture (the coefficient in this range is -1.061 (= -0.773 + -0.288). Thirdly, the size of the coefficient on GDP per capita within each of the four identified regions of GDPpc*culture is lower than that identified in regression 1. Indeed we can formally reject the hypothesis that the coefficient in any of these regions is equal to the coefficient on GDP per capita (–1.361) in model 1.[11] Increases in development do not yield as large reductions in corruption as that suggested by linear model. For comparison, a one standard deviation increase in GDP per capita now decreases corruption by between 1.2 and 1.7 depending on the culture and development regime in which the country currently exists. The effect of the same change in the linear model was 2.1 points. This is also consistent with the theoretical literature, where the strength of the relationship between corruption and development is generated in part by considering the full range of development at once.

Table 2: Corruption and Development Using Linear and Non-linear Models

Model	1	2	3	4
Dependent variable	corruption	corruption	corruption	corruption
Identifier (I) variable		$\log(GDPpc_{t-1}$ *Culture)		$\log(GDPpc_{t-1}$ *Culture)
$\log(GDPpc)_{t-1}$	-1.361 (26.55)**	-0.773 (6.31)**	-1.145 (16.66)**	-0.750 (5.66)**
$0.13 < I < 0.62$		-0.086 (2.17)*		-0.104 (1.89)
$0.62 < I < 0.76$		-0.173 (3.14)**		-0.199 (2.62)**
$I > 0.76$		-0.288 (4.95)**		-0.319 (3.69)**
CULTURE			-0.399 (4.61)**	0.067 (0.55)
Constant	16.183 (37.68)**	12.676 (17.21)**	16.100 (40.56)**	12.381 (13.33)**
Observations	325	325	325	325
R-squared	0.71	0.75	0.73	0.75

[11] The t-statistics for the range $I < 0.13$ is 22.99; for the range $0.13 < I < 0.62$ is 24.36; for the range $0.62 < I < 0.76$ is 23.98; and for the range $I > 0.76$ is 16.62.

Notes: **, * denotes significance at the 1% and 5% levels respectively. The dependent variable is corruption as measured by Transparency International. log(GDPpc) denotes the log of GDP per capita. Culture is measured using data from International Country Risk Guide. I denotes the variables used to identify thresholds between corruption and GDPpc.

The results can also be used to suggest other possible misspecifications in the previous empirical literature. In regression 3 we add to regression 1 the measure of culture that we had used to locate break points in the relationship between development and corruption. According to the results from this regression culture has a strong negative effect on corruption in the linear model. In regression 4 we then combine this new linear regression with the non-linear regression in 2, to test whether the effect of culture is direct (as in regression 3) or indirect (as in regression 2). We find in model 4 that while there is some reduction in the size of the effect of GDP per capita across the different regimes, the direct effect of culture variable is no longer statistically significant. Culture would therefore appear to affect corruption by generating non-linearities in the relationship between GDP per capita and corruption, remember the identifying variable *I* does not enter the regression directly, rather than having a direct effect itself.

This new empirical relationship between culture and development found here would suggest a specific agenda for further theoretical analysis of this point. Given that cultural variables have been found to be amongst those that are most robustly correlated with corruption (see for example Triesman, 2000) the potential sensitivity of other control variables is worth exploring. In Section 6 we use another variable often found to be correlated with corruption, openness to international trade and find a similar result.

Another way of considering the same question is by asking whether we prefer statistically model 2, where culture has an effect only by identifying non-linearities in the effect of GDPpc, or model 3, where its effect on corruption is direct. To test between these non-nested hypotheses we use a J-test (Davidson and MacKinnon, 1981). As stated above preference is established on the basis of whether the maintained model can explain the variation of the data of the competing model (Greene, 2003). The test results unambiguously favour model 2; where this result is independent of whether model 2 or model 3 is specified as the null or alternative hypothesis. The t-statistic with model 2 as the maintained hypothesis is 0.60, and 5.71 with model 3 as the maintained hypothesis (critical t-value = 1.95). That is, we unambiguously prefer the non-linear model, even though it has fewer explanatory variables on the right hand side of the regression.

6. Robustness

In this section of the paper we consider the robustness of the evidence found above to changes in the approach and alternative measures of the culture of a country. Firstly we consider whether the power of the interaction between culture and development in identifying the location of thresholds in the corruption-development relationship is derived from either the development or culture parts of the identifying variable.

It appears that it is not, although in both cases we find a smaller number of significant thresholds (see Table 3). Searching over the range of GDP per capita we find evidence of two significant thresholds. The first at the 70th percentile (GDPpc=$9,818; confidence interval 63rd-71st percentiles) and one at the 4th percentile (GDPpc=$237; confidence interval 3rd-4th percentiles).[12] Using culture as the identifying variable we again find evidence of two thresholds, the first at the 87th percentile (culture index = 5.65; confidence interval 87th-90th percentiles) and one at the 69th percentile (culture index = 4.78; confidence interval 10th-75th percentiles). Given the size of the confidence interval we do not search for a third threshold in this relationship.

In regressions 5 and 6 we report the results from a regression of GDPpc on corruption imposing the thresholds found using culture and GDPpc alone as the identifying variables, along with the J-test of whether we can prefer statistically each of these competing models to model 2. From the results of the J-test it is not clear which of the models we prefer relative to regression 2, we favour each time the alternative hypothesis of the competing model. GDPpc and culture would therefore both appear to be important partitioning variables and identifying differential effects of GDPpc on corruption and we cannot say, from a statistical perspective at least, which is more important.

As a final test on the culture variable we search across GDP per capita, and culture simultaneously to identify the location of thresholds. That is, we allow GDPpc and culture to compete to identify the location of any thresholds in the data. From this we find that the first two thresholds are identified better by GDP per capita rather than

[12] Evidence of a third threshold is also found (at the 63rd percentile) but it is statistically insignificant.

culture, the fit of the regression at the point of the threshold identified by GDPpc is better than that by culture.

Table 3: Corruption and Development Using Non-linear Models with Alternative Identifier Variables

Model	5	6	7	8	9	10
Dependent variable	Corruption	Corruption	Corruption	Corruption	Corruption	Corruption
Identifier (I) variable	GDPpc	Culture	Sequential GDPpc and Culture	Culture2* GDPpc	Openness* GDPpc	Openness* GDPpc
$\log(GDP)_{t-1}$	-1.454 (11.67)**	-1.082 (15.03)**	-1.347 (10.35)**	-0.877 (9.29)**	-1.432 (8.83)**	-1.051 (6.72)**
0.04>I<0.70	0.403 (6.33)**					
I>0.70	0.247 (3.16)**		0.197 (2.45)*			
0.67>I<0.87		-0.075 (2.80)**				
I>0.87		-0.183 (7.49)**				
$0.04>I_1<0.70$ $I_1 = \log(GDP_{t-1})$ $I_2 < 0.53$ $I_2 = \log(CULT)$			0.400 (6.32)**			
$0.04>I_1 <0.70$ $I_1 = \log(GDP_{t-1})$ $I_2 > 0.53$ $I_2 = \log(CULT)$			0.342 (5.11)**			
0.62<I<0.77				-0.080 (2.54)*		
I>0.77				-0.218 (6.63)**		
0.01>I<0.55					0.506 (4.54)**	0.418 (3.94**)
0.55>I<0.78					0.409 (3.47)**	0.267 (2.35**)
I>0.78					0.299 (2.46)*	0.099 (0.83)
$OPEN_{t-1}$						-0.001 (0.80)
Constant	14.254 (24.14)**	14.301 (26.80)**	13.699 (21.95)**	12.900 (19.32)**	13.524 (20.78)**	11.517 (17.84)**
Observations	325	325	325	325	325	282
R-squared	0.77	0.75	0.77	0.76	0.77	0.79
J-test (model 2) vs (model Z)	6.89	3.02	7.40	4.79	7.38	
J-test (model Z) vs (model 2)	3.16	3.04	2.74	0.47	4.88	

Notes: **, * denotes significance at the 1% and 5% levels respectively. The dependent variable is corruption as measured by Transparency International. log(GDPpc) denotes the log of GDP per capita. Culture is measured using data from International Country Risk Guide. I denotes the variables used to identify thresholds between corruption and GDPpc. The J-test test between model 2 in Table 2 with model Z, where Z is given as the model in the top row of the table.

Following from the results in model 5 these are known to be statistically significant. However, in contrast to model 5, where we found no more significant thresholds, using this sequential approach we find evidence of a third significant threshold. This threshold is identified using the culture variable and is located at the 54th percentile but has a very wide confidence interval. The results from this exercise are presented as model 7 in Table 3. Interestingly we can again not establish statistically whether this approach is preferred to that of model 2.

We next consider changes to the identifying variable X in equation 1. We first, expand the cultural variable to include measures of military and religion in politics and then secondly, the use of an alternative identifying variable in openness to international trade (in both cases these are interacted with GDPpc). For the latter Ades and Di Tella (1997, 1999) have previously shown that a higher degree of openness to international trade leads to lower corruption. They argue that more openness brings higher competition in the economy, which reduces corruption. Wei (2000), however, suggests that open countries experience greater losses from corruption than less open ones, because corruption creates distortions on foreign transactions. In a thought provoking survey, Winters (2004) suggests that trade policy can contribute positively to the fight against corruption.

For the expanded culture variable we find evidence of two rather than the three thresholds of before, and where the second of these thresholds has a very wide confidence interval. The first threshold is located at the 77th percentile of the GDPpc/culture distribution (confidence interval 76th – 77th percentiles) and the second at the 62nd percentile (confidence interval 1st – 76th percentiles). This is reported as model 8 in Table 3. An interesting result from this regression is the statistical preference for this expanded cultural variable; the J-test rejects the hypothesis that model 2 is preferred over model 8.

For openness to international trade three thresholds are found. The first at the 78th percentile of the GDPpc/openness distribution (confidence interval 1st – 76th percentiles), the second at the 55th the percentile (confidence interval 52nd –60th percentiles) and the third at the 1st percentile (confidence interval 1st – 52nd percentiles). Using openness as the identifier variable we can again answer the question of whether openness is best modelled as having a direct effect on corruption or an indirect effect. As with culture we again find evidence to suggest the effect is indirect. In model 10 we add to model 9 the openness variable, it has an insignificant relationship with corruption once we control for the non-linearities in GDP per capita and corruption.[13] In a regression without non-linearities (not reported) openness has a significant direct effect on corruption. Moreover a J-test suggests we prefer the model non-linearities (model 9) over the model with a direct effect of openness to international trade. The t-statistics are 6.81 and 0.80.

Unfortunately we have some limit to what we can say about favouring culture or openness as the identifying threshold variable. From the J-test we cannot establish statistical preference for model 9 with openness over model 2 which uses culture. Using the same sequential method as model 7 we can make a weak preference for openness. As before the first two thresholds are identified by GDP per capita but the third threshold is now identified by openness rather than culture. The third threshold is identified by openness at the 50th percentile of its distribution (confidence interval 31st – 70th percentiles). However a J-test of this model against model 7 suggests that we cannot statistically distinguish between the models however.

One plausible explanation of why we are unable to establish which of the non-linear models we prefer from the above analysis is that in each case the identifying variables place countries in similar regimes. In Tables 4 we try to assess the plausibility of this as an explanation using the three most distinct models from Sections 5 and 6, namely model 2, (GDPpc*culture as the identifying variable), model 8 (GDPpc*culture2 as the identifying variable) and model 9 (GDPpc*openness as the identifying variable). For each model we place observations in each of one of the three or four identified regimes (they lie above and below particular threshold points) counting from the lowest values of the identifying distribution upwards. We then cross-tabulate each of these models against each other in Tables 4 as parts a, b and c. If the models produce similar groupings of countries we would expect that most observations are placed in the same cells in alternative non-linear models. Approximately we expect that these should lie along the diagonal.

As shown by the tables this is generally the case, although there are enough exceptions to suggest that this is not a universal result. A good example of where the models produce similar results can be found in a comparison of Models 2 and 8 (Table 4a).

[13] Interestingly culture is significant when added to such a regression however.

Here of the 88 observations placed in the highest regime for model 8 (culture2*GDPpc > 77^{th} percentile), 86 of them appear in the highest regime for model 2 (culture*GDPpc > 76^{th} percentile).

Table 4a: Distribution of Countries across regimes under different non-linear models: Models 2 and 8

Model8/ Model2	Regime 1	Regime 2	Regime 3	*Total*	*Similarity Score*
Regime 1	34	0	0	*34*	*100*
Regime 2	145	6	0	*151*	*96.0*
Regime 3	7	41	2	*50*	*82.0*
Regime 4	0	4	86	*90*	*95.6*
Total	*186*	*51*	*88*	*325*	*94.2*
Similarity Score (%)	*78.0*	*80.4*	*97.7*	*83.7*	

Table 4b: Distribution of Countries across regimes under different non-linear models: Models 2 and 9

Model 9/ Model 2	Regime 1	Regime 2	Regime 3	Regime 4	*Total*	*Similarity Score*
Regime 1	2	30	0	0	*32*	*93.8*
Regime 2	2	101	27	8	*138*	*73.2*
Regime 3	0	6	21	13	*40*	*52.5*
Regime 4	0	0	17	55	*72*	*76.4*
Total	*4*	*137*	*65*	*76*	*282*	*73.4*
Similarity Score (%)	*50.0*	*73.7*	*41.5*	*72.4*	*65.6*	

Table 4c: Distribution of Countries across regimes under different non-linear models: Models 8 and 9

Model8/ Model9	Regime 1	Regime 2	Regime 3	*Total*	*Similarity Score*
Regime 1	4	0	0	*4*	*100*
Regime 2	131	6	0	*137*	*95.6*
Regime 3	29	22	14	*65*	*44.6*
Regime 4	7	13	56	*76*	*73.7*
Total	*171*	*41*	*70*	*282*	*78.0*
Similarity Score (%)	*76.6*	*53.7*	*80.0*	*74.1*	

Notes: For model 2 (GDPpc*culture) the four identified regimes are: regime 1 = I<0.13; regime 2 = 0.13>I<0.62; regime 3 = 0.62>I<0.76; regime 4 = I>0.76
For model 8 (GDPpc*culture2) the three identified regimes are: regime 1 = I<0.62; regime 2 = 0.62>I<0.77; regime 3 = I<0.77

For model 9 (GDPpc*openness) the three identified regimes are: regime 1 = I<0.01; regime 2 = 0.01>I<0.55; regime 3 = 0.55>I<0.78; regime 4 = I>0.78

At the other extreme, for the third regime in model 9 (55th percentile< GDPpc*openness >78th percentile) the observations are fairly evenly spread across regimes 2 to 4 in model 2 (Table 4b).

In order to get a better measure of similarity between models in the final row and column of each table we try to provide a summary statistic of how similar the different models are. For a given regime this is calculated as the 1 minus the percentage of observations in the most populous regime from the different model. So for example for regime 1 in model 8 the similarity index is calculated as [78={1 – 1*(186-145)/186}*100] i.e. of the 186 observations in regime 1 of model 8, 22 per cent are in regimes other than regime 2 of model 2. We perform a similar exercise for the model as a whole.

Without surprise the similarity of the models with culture as an identifying variable (models 2 and 8) with each other is higher than when openness is used as an identifying variable (model 9). According to the similarity measure, of the 325 observations in model 8 close to 84% are located in the most populous cells in model 2, whereas the corresponding number for model 2 against model 8 is 94%. The similarity index between models 2 and 9 is between 66 to 73 per cent and between 74 and 78 for models 8 and 9. Overall, it would therefore appear that to a reasonable degree the models produce similar results.

7. Conclusions

In this paper, motivated by recent developments in the theoretical literature on corruption, we provide empirical evidence of non-linearities in the relationship of this variable with development that statistically reject their linear relationship as shown by the existing literature. To perform this task we use the endogenous threshold methodology of Hansen (2000). Using different combinations of culture, development and openness to international trade to identify break-points in the regression coefficient for development we find between three and four corruption- development regimes that occur at different values of these identifying variables. While we are not able to establish statistical preference for any one of these identifying variables, and there is some tendency for each to place countries in similar regimes when we allow them to compete in terms of locating thresholds they each play a role. There would therefore appear to be some distinct information contained in each of the identifying variable as to why the relationship of development with corruption changes.

The results of the paper suggest that it may be important to test their future robustness in a number of dimensions. Firstly, their robustness against alternative forms of non-linearity (such as higher order terms of the independent variable) and alternative methods of locating threshold points (such as regression trees). Secondly, what other than openness and culture might be used to locate thresholds points. Finally, can we establish statistical preference for a particular form of non-linear model with a given identifying variable over another. Suggestions for this task may come from further development of theoretical models.

While this provides a clear direction of possible future empirical and theoretical work the paper might also be used to provide a strong rejection of the linear regression approach more commonly adopted and the progression of the research agenda on these lines. In a standard approach variables such as culture and openness are found to enter the regression equation with significant coefficients. However when we test this against the model in which their effect is in establishing the changes in the relationship between development and corruption, that is we do not enter them directly on the right hand side of the regression, we find clear statistical evidence in favour of this model. This means that, both culture and openness do not affect corruption directly; rather they affect the relationship between corruption and development. In all counts, we reject the linear model in favour of a non-linear one.

References

Ades, A. and R. Di Tella, 1997. The new economics of corruption: a survey and some new results. *Political Studies*, 45 (Special Issue), 496 – 515.

Ades, A. and R. Di Tella, 1999. Rents, competition and corruption. *American Economic Review*, 89, 982 – 993.

Andvig, J.C. and K.O. Moene, 1990. How corruption may corrupt. *Journal of Economic Behaviour and Organisations*, 13, 63 – 76.

Bardhan, P., 1997. Corruption and development: a review of issues. *Journal of Economic Literature*, 35, 1320 – 1346.

Blackburn, K., Bose, N. and Haque, M E., 2002. "Endogenous Corruption in Economic Development, Centre for Growth and Business Cycle Discussion Paper No. 22, December 2002, School of Economic Studies, University of Manchester.

Blackburn, K., Bose, N. and Haque, M E., 2006. "Incidence and Persistence of Corruption in Economic Development, *Journal of Economic Dynamics and Control*, 30, 2006: 2447-67.

Blackburn, K., Bose, N. and Haque, M E., 2005. "Public Expenditures, Bureaucratic Corruption and Economic Development, Centre for Growth and Business Cycle Discussion Paper No. 53, May 2005, School of Economic Studies, University of Manchester.

Cadot, O., 1987. Corruption as a gamble. *Journal of Public Economics*, 33, 223 – 244.

Chemlarova, V. and Papageorgiou, C., 2005, Nonlinearities in Capital-Skill Complementarity, *Journal of Econimic Growth* (forthcoming)

Davidson, R. and MacKinnon J. (1981) "Several Tests for Model Specification in the Presence of Alternative Hypothesis", *Econometrica*, 49, 781 – 793.

Del Monte, A. and Papagni, E. (2001): Public expenditure, corruption, and economic growth: the case of Italy. *European Journal of Political Economy,* 17 (1), 1 – 16.

Durlauf, S., Kourtellos, A and Minkin, A., 2001. The local Solow growth model, *European Economic Review*, 45, 928-940.

Dawid, Herbert, and Gustav Feichtinger, 1996, On the Persistence of Corruption, *Journal of Economics*, 64 (2), pp. 177–93.

Ehrlich, I. and F.T. Lui, 1999. Bureaucratic corruption and endogenous economic growth. Journal of Political Economy, 107, 270 – 293.

Fisman, R. and R. Gatti, 2002. Decentralisation and corruption: evidence across countries. Journal of Public Economics, 83, 325 – 345.

Greene, W. H. (2003). Econometric Analysis: 5th Edition, Prentice Hall, US.

Girma, S., Henry, M. Kneller R., and Milner C., 2003. Threshold and interaction effects in the openness-productivity growth relationship: The role of institutions and natural barriers, GEP Working paper 2003/32, The University of Nottingham.

Gyimah-Brempong, K., 2002. Corruption, Economic Growth and Income Inequality in Africa. *Economics of Governance*, 3, 183 – 209.

Hansen, B. E., 2000. Sample Splitting and Threshold Estimation, *Econometrica* , 68, 575 – 603.

Huang, P. H., and Wu, H., 1994. "More Order Without More Law: A Theory of Social Norms and Organizational Cultures," *Journal of Law, Economics, and Organization*, 10 (2), 390 – 406.

Huntington, S. P., 1968. Political Order in Changing Societies. Yale University Press, New Haven.

Jain, A.K. (ed.), 1998. The Economics of Corruption. Kluwer Academic Publishers, Massachusettes.

Johnson P. and Takeyama, L., 2001. Initial conditions and economic growth in the US states, *European Economic Review*. 45, 919 – 927.

Knack, S. and P. Keefer, 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics*, 7 (3), 207 – 227.

La Porta R., Lopez-de-Silanes F., Shlieifer, A., and Vishny, R.W., 1999. The Quality of Government, *The Journal of Law, Economics and Organiszation*, 15, 222-279.

Leff, N. H., 1964. Economic Development through Bureaucratic Corruptoin. In A. K. Jain (ed.), *The Economics of Corruption*, Kluwer Academic Publishers, Massachusettes.

Leys, C., 1970. What is the Problem about Corruption? In A. J. Heidenheimer (ed.), *Political Corruption: Reading in Comparative Analysis,* Holt Reinehart, New York.

Li, H., Xu L.C., and Zou, H., 2000. Corruption, Income Distribution and Growth, *Economics and Politics*, 12, 155-182.

Lui, F., 1985. An Equilibrium Queuing Model of Corruption. *Journal of Political Economy*, 93, 760 – 781.

Massanajala, W. and Papageorgiou C., 2004. The Solow model with CES technology: Nonlinearities and parameter heterogeneity, *Journal of Applied Econometrics*, 107, 407 – 437.

Mauro, P., 1995. Corruption and growth. *Quarterly Journal of Economics*, 110, 681 – 712.

Mauro, P., 2004. Persistence of Corruption and Slow Economic Growth. *IMF Staff Papers*, 51 (1), 1 – 18.

Montinola, G.R. and R.W. Jackman, 1999. Sources of corruption: a cross-country study. British Journal of Political Studies, 32, 147 – 170.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, 1993. 1993, "Why Is Rent-Seeking So Costly to Growth?" *American Economic Review, Papers and Proceedings*, 83 (2), 409–14.

Paldam, M. 2002. The big pattern of corruption, economics, culture and seesaw dynamics. European Journal of Political Economy, 18, 215 – 240.

Papageorgiou, C., 2002 Trade as a threshold variable for multiple regimes, *Economics Letters*, 77, 85 – 91.

Persson, T., Tabellini, G., and Trebbi, F., 2001. Electoral Rules and Corruption, NBER Working Paper 8154.

Putnam, R.D., 1993. Making Democracy Work: Civic Traditions in Modern Italy, (Princeton, New Jersey: Princeton University of Press.

Rauch, J.E. and Evans, P.B., 2000. Bureaucratic structure and bureaucratic performance in less developed countries. Journal of Public Economics, 76, 49 – 71.

Sah, R.K., 1991. Social Osmosis and Patterns of Crime. *Journal of Political Economy*, 99, 1272 – 1295.

Sarte, P.-D., 2000. Informality and rent-seeking bureaucracies in a model of long-run growth. Journal of Monetary Economics, 46, 173 – 197.

Tanzi, V. and H. Davoodi, 1997. Corruption, public investment and growth. *IMF Working Paper* No.WP/97/139.

Tirole, J. 1996. A Theory of Collective Reputation (with applications to the persistence of corruption and to farm quality). *Review of Economic Studies*, 63, 1 – 22.

Treisman, D., 2000. The causes of corruption: a cross-national study. *Journal of Public Economics*, 76, 399 – 457.

Wei, S., 2000. How taxing is corruption on international investors? Review of Economics and Statistics, 82, 1 – 11.

Winters, L.A., 2004. 'Trade liberalisation and economic performance: an overview', *The Economic Journal*, 114 (493), F4 – F21.