# Does E-government Reduce Corruption?*

Thomas Barnebeck Andersen

Department of Economics, University of Copenhagen

October 2007

**Abstract**

Countries that have invested more in e-government have also seen larger reductions in levels of corruption. The statistical association between e-government and corruption is economically larger than the comparable association between real GDP per capita growth and corruption. Compared to OLS, 2SLS doubles the impact of e-government on corruption but the difference is not statistically significant. In conclusion, results suggest that e-government can make important headway in the fight to reduce corruption.

Keywords: Corruption, ICT, E-Government

JEL Classifications: D73, H11, O1, O57

*I would like to thank Phil Abbott, Carl-Johan Dalgaard, Henrik Hansen, Theo Ib Larsen, Martin Rama, John Rand, Pablo Selaya, and Clay G. Wescott for helpful comments, which enabled me to write this paper. All remaining errors and omissions are mine. Address for correspondence: Thomas Barnebeck Andersen, Department of Economics, University of Copenhagen, Studiestræde 6, DK-1455, Copenhagen K, Denmark; Email: Thomas.Barnebeck.Andersen@econ.ku.dk

"E-government offers a partial solution to the multifaceted problem of corruption. It reduces discretion, thereby curbing some opportunities for arbitrary action. It increases chances of exposure by maintaining detailed data on transactions, making it possible to track and link the corrupt with their wrongful acts. By making rules simpler and more transparent, e-government emboldens citizens and businesses to question unreasonable procedures and their arbitrary application." (Global Corruption Report 2003, p. 30)

# 1 Introduction

Corruption is commonly considered to be one of the most significant impediments to economic development.[1] In fact, commentators and NGOs have emphasized that in order to meet the Millennium Development Goals corruption must be reduced.[2] In this regard, the result presented here, namely that an increase in the use of e-government[3] is likely to reduce corruption, is constructive. The mechanisms through which e-government works are straightforward: E-government reduces contact between corrupt officials and citizens and increases transparency and accountability.

As the opening quotation demonstrates, the potential of e-government in the fight against corruption has not slipped the attention of practitioners. The Asian Development Bank (ADB) have provided a long list of examples of e-government initiatives worldwide along with interesting anecdotal evidence intended to document achievements (Wescott 2003). In Pakistan, the entire tax department is undergoing restructuring and information and communication technology (ICT) systems are being introduced with the stated purpose of reducing contact between tax collectors and tax payers. In the Philippines, the Department of Budget and Management has established an on-line e-procurement system that allows public bidding for suppliers. This system has increased transparency in transactions.

[1] A classic reference is Mauro (1995), who finds that corruption hampers economic growth. For two good complementary reviews of the literature on corruption, see Bardhan (1997) and Svensson (2005).

[2] In a press release (dated 14 September 2005), Transparency Internation claims that "Millennium Development Goals are unreachable without commitment to fighting corruption".

[3] One definition of e-government (or digital government) is "public sector use of the Internet and other digital devices to deliver services, information, and democracy itself." (West 2005, p. 1.) Another definition is that e-government is "the process of connecting citizens digitally to their government in order that they might access information and services offered by government agencies." (Lau et al. 2007, p. 2.)

In South Korea, the Online Procedures Enhancement for Civil Applications allows ordinary citizens to monitor applications for permits or approvals where corruption is most likely to take place; it also allows questions to be raised in case irregularities are detected. In the Indian state of Andhra Pradesh, where 40% of its 76 million people cannot read, 214 deed registration offices have been fully computerized. This has made the process of deed registration easy and transparent. The process started in April 1998 and by February 2000 about 700,000 documents had been registered. Before the introduction of online registration, opaqueness of procedures forced citizens to employ middlemen who used corrupt practices to obtain services. In several Asian countries, governments are introducing smart cards that help citizens access health-care services without having to provide corruption-prone cash payments for these services.

An impressive and well-known example of the potential of e-government in empowering citizens to challenge corrupt and arbitrary bureaucratic action is the Bhoomi (meaning land) system from Karnataka, India, where the introduction of an electronic land record system serving roughly 7 million farmers has saved clients some 1.32 million work days in waiting time and Rs. 806 million in bribes (World Bank 2004).[4] The main function of the Bhoomi system is to maintain records of rights, tenancy and cultivation, which are crucial for transferring or inheriting land and obtaining loans. Under the old system, some 9,000 village accountants, each serving three or four villages, maintained land records. Farmers had to seek out a village accountant in order to obtain a copy of the record or make changes. Accountants were not easily accessible and farmers faced long delays; two out of three paid bribes, and over two-thirds paid more than Rs. 100, compared to the official service fee of Rs. 2. Under the electronic Bhoomi system, farmers can enter a Bhoomi kiosk and get these records or file for changes in 5-30 minutes. Moreover, all requests are served on a first-come, first-served basis.

Other examples of e-government include Christal in Argentina, a website aiming at disseminating online information concerning the use of public funds; the Central Vigilance

[4]See Chawla and Bhatnager (2004) for a case study of the Bhoomi system.

Commission website in India, where the public among other things can report information about wrongdoings of public servants; an on-line Customs Bureau system in the Philippines, which has lessened the cost of trade for businesses, reduced opportunities for fraud and boosted revenue collection of the Customs Bureau; and several computerized interstate check posts in Gujarat, India, which has significantly reduced corruption at check posts.[5]

Summing up, the anecdotal evidence suggests that e-government eliminates many opportunities for corruption. However, this proposition has not been subjected to systematic empirical scrutiny.[6] The present paper attempts to correct this shortcoming. In doing so, the paper also proposes a novel identification strategy, which should be of some worth to empirical (cross-country) researchers interested in ICT.

The discussion proceeds as follows: Section 2 provides details on specification and identification issues, Section 3 discusses the data, Section 4 provides econometric results and Section 5 concludes.

# 2 Empirical Framework

## 2.1 Model specification

Since governance indicators are somewhat persistent, empirical papers studying the determinants of corruption usually rely on the variation in corruption levels across countries (between-country variation). A long time-span is needed in order to observe large changes in corruption levels within countries (within-country variation). For present purposes, however, the choice between within-country and between-country variation is straightforward. The paper's concern is whether e-government has had a measurable impact on corruption; and since e-government is a fairly new technology, it only makes sense to study whether changes in e-government can explain changes in corruption over the time-span in which e-

[5] See <http://www1.worldbank.org/publicsector/egov/anticorruption&t.htm> for more information on these and other initiatives.

[6] The lack of hard evidence linking e-government and corruption is fully recognized by a leading e-government proponent (United Nations Development Programme, APDIP e-NOTE 8, 2006).

government has actually been in operation. Focus must therefore be on the within-country variation in corruption levels.

Consequently, I rely on a model which attempts to explain changes in corruption, $DCCI_i = CCI_{2006,i} - CCI_{initial,i}$, by changes in e-government, $DEGOV_i = EGOV_{2006,i} - EGOV_{initial,i}$, and the initial level of corruption, $CCI_{initial,i}$. The following specification is a natural starting point:

$$DCCI_i = \alpha_0 + \alpha_1 DEGOV_i + \alpha_2 CCI_{initial,i} + \varepsilon_i. \tag{1}$$

The inclusion of the initial level of corruption means that (1) is equivalent to the following levels regression with a lagged dependent variable:

$$CCI_i = \alpha_0 + \alpha_1 DEGOV_i + (\alpha_2 + 1)\, CCI_{initial,i} + \varepsilon_i. \tag{2}$$

Importantly, inclusion of a lagged dependent variable in a cross-section regression such as (2) is a simple way to account for historical factors that cause current differences in levels of corruption, but which are difficult to account for otherwise (Wooldridge 2000). Countries with historically high levels of corruption are perhaps less likely to implement e-government, and this would render $\text{Cov}(DEGOV, \varepsilon) \neq 0$ had $CCI_{initial}$ not been controlled for. Thus, relying on the specification in (1) significantly reduces any potential omitted variables endogeneity problem. In addition, it captures the fact that NGOs, multilateral donors as well as bilateral donors have directed some attention to fighting corruption in recipient countries, implying that we should probably expect more improvement at the bottom end of the corruption distribution.

There may still be reason to suspect potential endogeneity problems. For instance, e-government could be part of a wider public reform package. A positive (partial) correlation between increases in e-government and improvements in the level of corruption would be spurious if these other dimensions of the reform program were in fact causing the reduction in corruption. To the extent that public reform programmes lead to higher growth rates in

GDP per capita, this omitted variables problem can be remedied by including the growth rate in real GDP per capita over the period, $GYCAP_i$, as an additional control in (1). This leaves us with the following specification:

$$DCCI_i = \beta_0 + \beta_1 DEGOV_i + \beta_2 CCI_{initial,i} + \beta_3 GYCAP_i + \zeta_i. \tag{3}$$

If there are feedback effects between changes in corruption and real GDP per capita growth, including $GYCAP$ is not neccesarily a good strategy. However, if $\alpha_1$ and $\beta_1$ are equal, i.e. if the inclusion of $GYCAP$ has no (statistically significant) effect on the $DEGOV$ slope estimate, one can take this as an indication (not a formal test) that $\alpha_1$ is not influenced by deliberate policy changes.

## 2.2 Identification strategy

OLS informs us of partial correlations; whether it informs us of causal effects is less clear. A causal interpretation of the association between e-government and corruption requires pure exogenous variation in e-government. This also addresses the concern that e-government is measured with error, which may cause attenuation bias in OLS.

The identification strategy exploits the fact that climate has non-trivial implications for modern information and communication technology and power distribution. Computers generally prefer moderate (relative) humidity as opposed to extreme humidity.[7] A rainforest, for instance, provides a bad working environment for a computer. Under such extreme conditions, computers may simply short-circuit with various damaging consequences for system components.[8] Under less extreme conditions, humidity may lead to corrosion and

[7] *Absolute humidity* is the mass of water vapor divided by the mass of dry air in a volume of air at a given temperature. The hotter the air is, the more water it can contain. *Relative humidity* is the ratio of the current absolute humidity to the highest possible absolute humidity (which depends on the current air temperature). A reading of 100 percent relative humidity means that the air is totally saturated with water vapor and cannot hold any more.

[8] The Los Angeles Times (March 15, 1999) provides an "everyday life" illustration from Peru's Amazon River basin of just how bad computers and tropical humidity mix ("Up the River With Heat, Humidity and Computer" by Lawrence J. Magid).

possible condensation risk, which can also damage equipment.[9] On top, humidity makes cooling the computer more difficult, and a high temperature may cause premature aging and failure of chips.

Humidity concerns are likely to be minor. A much more important climate-related concern is thunderstorms.[10] According to the National Lightning Safety Institute, lightning in the United States accounted for 101,000 laptop and desktop computer losses in 1997.[11] More specifically, computer chips based on solid-state electronics[12] are extremely vulnerable to (cloud-to-ground) lightning if left unprotected.[13] Brief overvoltages caused by lightning can immediately destroy solid-state components or weaken them to the point that they fail some months after the lightning event.[14] The problem is particularly acute for multi-port appliances, i.e. appliances connected to several different systems (IEEE 2005). A modem connected to the telephone line is also connected to a computer, which in turn is connected to grid power. This leaves a modem exposed to voltage surges. Moreover, during a lightning strike, high voltages can enter the computer through a phone line connected to the modem.

In general, lightning discharges can enter electronic equipment inside a residence in four principal ways (IEEE 2005). First, lightning can strike the network of power, phone and cable television wiring. This network, particularly when elevated, acts as an effective collector of lightning surges. The wiring then conducts the surges directly into the residence, and then to the connected equipment. In fact, the initial lightning impulse is so strong that

[9] A high level of humidity causes internal components to rust and degrade essential properties such as electrical resistance or thermal conductivity.

[10] In fact, lightning research first became particularly active in the late 1960's because of the danger of lightning to aerospace vehicles and solid-state electronics used in computers and other electronic devices (NASA 2007).

[11] <http://www.lightningsafety.com/nlsi_lls/nlsi_annual_usa_losses.htm>

[12] Solid-state electronics refers to an electronic device in which electricity flows through solid semiconductor crystals rather than through vacuum tubes. Transistors, made of one or more semiconductors, are at the heart of modern solid-state devices. In the case of integrated circuits, millions of transistors can be involved. Microprocessors are the most complicated integrated circuits. They are composed of millions of transistors that have been configured as thousands of individual digital circuits, each of which performs some specific logic function (see Kressel 2007 for an enjoyable discussion with a historical perspective).

[13] The New York Times (April 6, 2000) provides an "everyday life" illustration of the damaging consequences of lightning for electronic devices ("The High Cost of Underestimating Lightning" by Lynn Ermann).

[14] Science and Technology Review, May 1996, "Mitigating Lightning Hazards". Available online at: <http://www.llnl.gov/str/05.96.html>.

equipment connected to cables several miles from the site of the strike can be damaged. Second, when lightning strikes directly to or nearby air conditioners, satellite dishes, exterior lights, etc., the wiring of these devices can carry surges into the residence. Third, lightning may strike nearby objects such as trees, flagpoles, road signs, etc., which are not directly connected to the residence. When this happens, the lightning strike radiates a strong electromagnetic field, which can be picked up by the wiring in the building, producing large voltages that can damage equipment. Finally, lightning can strike directly into the structure of the building. This type of strike is extremely rare, even in areas with a high lightning density (flashes per unit area per unit time).

Lightning also causes damage to power infrastructure. The probability of lightning-caused power interruptions or equipment damage scales linearly with the lightning density. In the United States, lightning is estimated to be the direct cause of one-third of power quality disturbances (Chisholm and Cummins 2006).

It is possible to take measures to protect equipment. An air-conditioned humidity-controlled room takes care of heat and humidity risks, whereas a high-quality surge protector provides protection against voltage spikes. But the crux of the matter is that if one lives in a hot, humid environment with a high annual lightning density, this adds to the costs of using modern electronics, including a computer.

The warmest and most humid places on Earth are generally located closer to the equator. In addition, thunderstorms occur most often in the tropical latitudes over land, where the air is most likely to heat quickly and form strong updrafts (Encyclopædia Britannica 2007). Many tropical land-based locations experience over 100 thunderstorm days per year. More specifically, only regions located between 35 degrees South latitude and 35 degrees North latitude experience more than 50 thunderstorm days per year on average. In places between 23.5 degrees South latitude and 23.5 degrees North latitude, called the geographical tropics, the average is even higher as they can experience more than 100 thunderstorm days per year. Strikingly, in some places between 10 degrees South latitude and 10 degrees North

latitude, one may experience more than 180 thunderstorm days per year.[15] Outside the tropics, thunderstorms are less frequent and more seasonal, occurring in those months where heating is most intense. For instance, in East and Central Europe the average is between 20 to 40 thunderstorm days per year, whereas Northern Europe experiences 5 to 20 yearly thunderstorm days (Encyclopædia Britannica 2007).

In sum, essential technologies for e-government such as the computer are less adapted to the hot, humid and lightning prone tropical climate and hence more costly to adopt in this environment. In addition, a large number of thunderstorm days is likely to cause frequent power outages, which in itself is a major nuisance for user of electronic equipment. As a consequence, modern information and communication technologies are ceteris paribus likely to spread more slowly in areas with tropical climate. As instrument for e-government, I therefore use the *percentage land area in the tropics*.[16]

Tropical land area is certainly exogenous in a deep sense. However, this does not ensure validity of the exclusion restriction. Validity rests on a redundancy condition in (1). Climate-related circumstances are likely to map into levels of corruption. For instance, to the extent that countries in the tropics have a higher endowment of natural resources, this is likely to affect levels of corruption (resource curse). However, this is controlled for with the inclusion of $CCI_{initial.}$ As I will outline in more detail below, the period under study

[15] It has been estimated that at any one moment there are roughly 1,800 thunderstorms in progress throughout the world (Encyclopædia Britannica 2007).

[16] It would be preferable to use direct climatic measures. Unfortunately, cloud-to-ground lightning data with a high spatial and temporal accuracy are not available. More specifically, the pertinent characteristic of lightning in the evaluation of risk to electronic equipment and electric power systems is the *ground flash density*, expressed as the number (per unit area per unit time) of cloud-to-ground lightning flashes. Since the mid-1980's, it has been possible to measure ground flash density more directly using networks of electromagnetic sensors. Lightning Location Systems (LLS) are able to resolve individual ground strikes comprising a flash with high spatial and temporal accuracy. Unfortunately, many parts of the world, particularly the developing world, are not covered by the LLS data. Optical lightning observations (from a low-earth orbit satellite) produced by NASA's Optical Transient Detector and Lightning Imaging Sensor do exist. These are global observations of *total* lightning, i.e. both intra-cloud and cloud-to-ground lightning. However, these data do not separate out cloud-to-ground and intra-cloud lightning incidences (Chisholm and Cumming 2006). Over mid-Northern latitudes, for instance, it is known that some 80% of lightning occurs within clouds (intra-cloud). About 20% of all lightning is cloud-to-ground, while an extremely small percentage is cloud-to-sky and cloud-to-cloud (Science and Technology Review, May 1996, "Mitigating Lightning Hazards").

is 1996-2006. Validity of the exclusion restriction therefore requires that the share of the land area in the tropics has no impact on *changes* in corruption over the 1996-2006 period once I control for the *level* of corruption in 1996, $CCI_{1996}$. Formally, let $z$ denote the share of tropical land area, validity then requires $\text{Cov}(\varepsilon, z) = 0$, where $\varepsilon$ is the error term in (1). This remains a maintained assumption, although I will discuss over-identification (OID) issues in Section 4.2.

# 3 Data

## 3.1 Corruption

In order to measure corruption, I rely on the well-known Control of Corruption Index ($CCI$) compiled by Kaufmann et al. (2007). The $CCI$ measure, which ranges from $-2.5$ (worst) to 2.5 (best), is available biannually from 1996 to 2002 and then annually from 2002 onwards. The $CCI$ indicator attempts to measure "the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture by elites and private interests" (ibid. p. 4). The indicator is based on a large number of individual data sources, which are then aggregated into one measure by an unobserved components model. This means that the aggregate measure is a weighted average of the underlying individual data sources, with weights reflecting the precision of each of these underlying data sources. By virtue of being a solution to a statistical signal extraction problem, the aggregate $CCI$ indicator is presumably more informative than any individual data source.[17] This makes the $CCI$ is the most comprehensive measure of corruption around.[18]

[17] Svensson (2005), however, notes that the aggregation procedures used by Kaufmann et al. presumes that subindicator measurement errors are independent across sources. Yet, in reality, errors are likely to be highly correlated, because the producers of the different indices read the same reports and most likely gauge each other's evaluations. If this independence assumption is relaxed, the gain from aggregating a number of different reports is less clear.

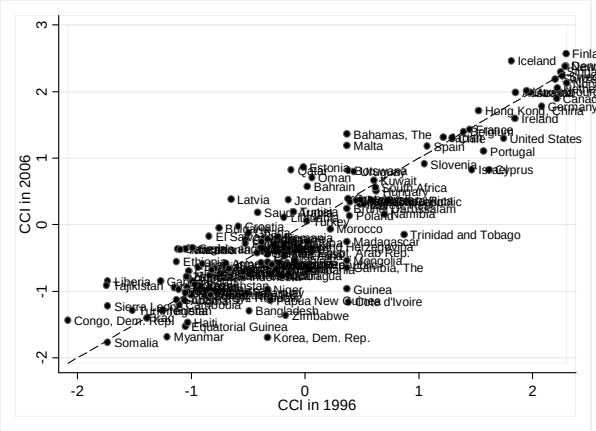
[18] The widely reported Corruption Perception Index ($CPI$) compiled by Transparency International forms part of the CCI measure (see Kaufmann et al. 2007, Table A13). Reasuringly, however, the simple correlation between $CCI$ and $CPI$ is 0.97. A high correlation is not unexpected since corruption reflects an underlying institutional framework (Svensson 2005).

$DCCI$ is calculated as the difference between $CCI$ in 2006 and 1996, i.e.

$$DCCI_i = CCI_{2006,i} - CCI_{1996,i}.$$

As noted by Kaufmann et al. (2007), despite being somewhat persistent governance indicators do change over relatively short periods such as a decade. This is illustrated in figure 1, which plots the 1996 $CCI$ score on the horizontal axis, the 2006 $CCI$ score on the vertical axis and a 45-degree line. Countries above the 45-degree line corresponds to improvements in corruptions, while countries below the line saw deteriorations in corruption.

**Figure 1:** Within-country variation in corruption levels over the 1996-2006 period.




Notes: Scatter plot of Control of Corruption in 1996 (horizontal axis) versus Control of Corruption in 2006 (vertical axis). The full line is the 45-degree line. The sample used is the largest estimation sample used in the empirical analysis below (number of obs. = 149).

## 3.2 E-government

The e-government variable, $EGOV$, used in this paper ranges from 0 (low) to 100 (high). The variable was compiled by a research team headed by Darrell M. West of Brown University during June and July 2006.[19] The methodological framework, upon which $EGOV$ is based,

[19] The report is available online at: <http://www.insidepolitics.org/egovt06int.pdf>.

is outlined in a book published on Princeton University Press (West 2005). Peer-acceptance is thus an important trait of $EGOV$.[20]

The team made an assessment of $1,782$ national government websites for 198 nations around the world. A range of sites within each country were analyzed to get a full sense of what is available in particular nations. Among sites analyzed where those of executive offices (president, prime minister, ruler, party leader, or royalty), legislative offices (Congress, Parliament, or People's Assemblies, etc.), judicial offices (such as major national courts), cabinet offices, and major agencies serving crucial functions of government (including health, human services, taxation, education, interior, economic development, administration, natural resources, foreign affairs, foreign investment, transportation, military, tourism, and business regulation). Websites were evaluated for the presence of various features dealing with information availability, service delivery, and public access. Features assessed included having online publications, online database, audio clips, video clips, non-native languages or foreign language translation, commercial advertising, premium fees, user payments, disability access, privacy policy, security features, presence of online services, number of different services, digital signatures, credit card payments, email address, comment form, automatic email updates, website personalization, personal digital assistant (PDA) access, and an English version of the website. Importantly, a common service featuring on government websites in the 57 countries having services that were fully executable online in 2006 was the possibility of reporting fraud and corruption (West 2006). Moreover, in 91% of all countries government websites offered visitors email contact material so that visitors could email a person in a government department other than merely the webmaster. This reduces the distance between officials and citizens.

There is no data on e-government dating back to 1996. Nevertheless, the technology is of a fairly recent origin. E-government relying on the World Wide Web (the Web) cannot

[20] In their Global E-Government Readiness Report 2005, the United Nations has also compiled an e-government variable. The United Nations variable and the e-government variable used in the present paper have a rank correlation of almost 0.7. I have chosen against the United Nations variable because it, unlike $EGOV$, has not been peer-reviewed. Reassuringly, however, the conclusions reported in this paper carries through using this alternative e-government variable.

be older than the Web itself, which is dated back to 1991 (West 2005). However, the process of commercializing the Web took off with the release of the Netscape browser in December 1994. Since $EGOV$ only measures e-government relying on the Web, e-government in 1996 is coded as zero. Consequently, $DEGOV$ is calculated as

$$DEGOV_i = EGOV_{2006,i} - EGOV_{1996,i} = EGOV_{2006,i}.$$

The $EGOV$ variable only measures a subset of e-government, namely "Internet-based e-government". The use of smart cards in healthcare, for example, is not directly captured by $EGOV$. However, it seems plausible that Internet-based e-government and other types of e-government are highly (positively) correlated. At any rate, the Internet remains the most popular e-government delivery system (West 2005).[21]

### 3.3 Additional variables

Real GDP per capita growth, $GYCAP$, is calculated using World Development Indicators (WDI) 2007. That is,

$$GYCAP_i = \ln\left(GDPCAP_{2005,i}\right) - \ln\left(GDPCAP_{1996,i}\right).$$

Real GDP per capita in 2005 is used since this is the last year available in the WDI 2007.

The WDI regional classification is used to create regional dummies, which are employed in the robustness analysis in section 4.1.1. These are the following: East Asia & Pacific (eap), Europe & Central Asia (eca), Latin America & the Caribbean (lac), Middle East & North Africa (mena), South Asia (soa), Sub-Saharan Africa (ssa) and North America (na).

I rely on two measures of land area in the tropics, *geographical* and *ecological tropics*. The geographical tropics are defined as the region of the Earth in which the Sun passes directly overhead at some point during the year. This includes the area between 23.5 degrees

[21] In the United States, 81% of all federal e-government initiatives are delivered via the Web (West 2005). In Britain, Directgov - an official Webpage launced in 2004 - aims to contain the whole of the British state in one place: <http://www.direct.gov.uk/en/index.htm>.

North latitude (Tropic of Cancer) and 23.5 degrees South latitude (Tropic of Capricorn). Ecological tropics in turn are based on the so-called Koeppen-Geiger climate classification system. The three tropical zones are tropical rainforest climate ($Af$), tropical wet and dry or savanna climate ($Aw$) and tropical monsoon climate ($Am$).[22] Data is from the Center for International Development at Harvard University and used in Sachs (2000).[23] I denote the share of the land area in the geographical tropics (respectively ecological tropics) $GEOTROPICS$ (respectively $ECOTROPICS$). Summary statistics and correlations of all variables are provided in table 1.

- Table 1 about here -

# 4 Econometric Results

## 4.1 OLS and LAD regressions

Table 2 shows regression results. Panel A reports results from OLS. The point estimate for $DEGOV$ in column (1), where the only additional control is $CCI_{1996}$, is 0.021. With a standard error of 0.008, the point estimate is more than 2.6 standard errors above zero. Following standard usage of letting a ratio of coefficient estimate to standard error ($t$-ratio) of 1.96 (respectively 1.65) indicate statistical (respectively marginal statistical) significance, the coefficient is estimated with sufficient precision to be regarded as larger than zero.

The coefficient estimates associated with $CCI_{1996}$, which are always negative, have (absolute) $t$-ratios that are never below 3.8 in table 2, panel A. Countries with low levels of corruption in 1996 (i.e. a high $CCI_{1996}$) on average saw less improvement in corruption over the subsequent decade. As mentioned in Section 2, NGOs, multilateral donors as well as bilateral donors have focused on fighting corruption in recipient countries, so there has been considerable attention directed towards the problem at the bottom end of the corruption

[22] A map associated with the classification system can be viewed <http://koeppen-geiger.vu-wien.ac.at/index.htm#maps>.

[23] Data are available online at: <http://www.cid.harvard.edu/ciddata/geographydata.htm#general>.

distribution.

The coefficient estimate associated with $GYCAP$ is positive and has a $t$-ratio of 2.18 ($\simeq 0.435/0.2$) in table 2, panel A. Higher real GDP per capita growth is on average associated with larger reductions in corruption. This may be due to an indirect effect if GDP per capita growth is the result of wider public reform. Alternatively, it could be due to a direct effect if GDP per capita growth alters the relative attractiveness of pursuing legal activities vis-a-vis corrupt activities. Finally, there could be reverse feedback effects in operation.

As a first indirect test of whether $DEGOV$ is driven by omitted factors, I conduct a $t$-test of $H0 : \beta_1 = \alpha_1$ using the same estimation sample, i.e. columns (2) and (3) in panel A. The test yields a $p$-value of 0.68. I take this as indicating that the partial correlation between changes in e-government and changes in corruption, where the initial level of corruption is partialled out, is not driven by omitted factors.

This rather parsimonious specification can explain about one-fifth to one-fourth of the total variation in changes in corruption over the decade.[24] This should be seen against the background of (presumably) some measurement error in both the dependent variable and the e-government variable.

- Table 2 about here -

As a first robustness check, panel B reports median (or least absolute deviations, LAD) regressions with bootstrapped standard errors. The coefficients and standard errors for $DEGOV$ are almost identical to the OLS results, indicating the the association is robust. Robustness is pursued in more detail in Section 4.1.1, but first I will address economic significance.

Using the OLS coefficients from table 2, table 3 reports standardized coefficients. The most conservative estimate of e-government's impact on corruption is obtained using column (1), which is also based on the largest estimation sample. The effect on $DCCI$ of a one standard-deviation increase in $DEGOV$ is an increase of 0.282 standard-deviation

[24] The levels version of equation (3), i.e. equation (2) with $GYCAP$ included, accounts for roughly 85% of the variation in the *level* of corruption in 2006.

units. Column 2 allows for a comparison of the *size* of the association between $DEGOV$, $GYCAP$, and $DCCI$. Whereas a one standard-deviation change in $DEGOV$ increases $DCCI$ by 0.294, the comparable effect on $DCCI$ of changes in $GYCAP$ is 0.212. Hence, the association between e-government and improvements in corruption is economically larger than the association between real GDP per capita growth and improvements in corruption.

- Table 3 about here -

What does the OLS-based economic significance imply for the "average" sample country? Consider Jamaica, which is the average country in the estimation sample in terms of corruption and e-government.[25] The most conservative estimate is obtained using column (1) in table 3. In the estimation sample, a one standard-deviation increase in $DEGOV$ would move Jamaica from $CCI$ country-ranking 76 to country-ranking 64 in 2006.

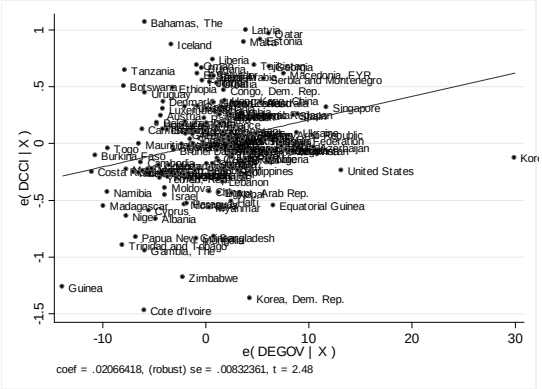
### 4.1.1 Additional OLS issues

A first check of robustness of the partial correlation was provided in table 2, panel B, where I reported median regressions with bootstrapped standard errors. In this section, I report further checks of the robustness of the partial correlation between changes in e-government and changes in corruption.

Visually, robustness is best assessed by inspecting the partial regression leverage plot associated with the full sample (Krasker et al. 1983). This plot, which is shown in figure 2, shows that no individual country (or cluster of countries) seems to be driving the positive partial correlation between $DEGOV$ and $DCCI$.

[25] It has e-government and corruption "closest" to median values in the estimation sample with 149 countries.

**Figure 2:** Association between changes in e-government and changes in corruption when initial corruption level and real per capita growth are partialled out.




Notes: e(DEGOV | X) is the residuals of an OLS regression of DEGOV on a constant, CCI in 1996 and GYCAP. e(DCCI | X) is the residuals from a regression of the DCCI on a constant, CCI in 1996 and GYCAP. The plot is then constructed as a scatter plot of the two vectors of residuals. The full line is the associated simple regression line.

Another simple approach to assess robustness is to exclude one region at a time, using the WDI classification, and then re-estimate column (1) in table 2, panel A to see what this does to the *DEGOV* slope estimate. Results are provided in table 4.

- Table 4 about here -

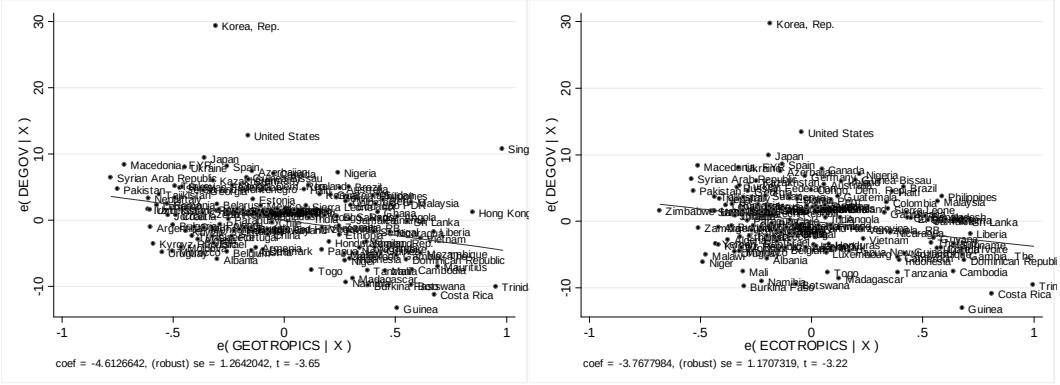
Inspection of the table shows that the partial correlation is robust and stable. The only time that the ratio of coefficient estimate to standard error is below 1.97 is when Sub-Saharan Africa is excluded and when Europe & Central Asia is excluded. In these cases, however, $t$-values remain at or above marginal significance. In addition, in all columns coefficient estimates are not statistically different from 0.02, the same as in table 2!

## 4.2 2SLS regressions

This section turns to the endogeneity issue. A natural starting point is to gauge visually the quality of instruments. Figure 3 plots the partial association between *DEGOV* and the share of land area in the tropics when the influence of past corruption levels and growth

in real GDP per capita are partialled out. This is the partial correlation used to obtain identification. In full accord with the hypothesis underlying the identification strategy, countries with a larger share of their land area in the tropics (geographical tropics in panel A and ecological tropics in panel B) have less e-government, after controlling for differences in past corruption levels and growth in real GDP per capita.

**Figure 3:** Association between e-government and share of tropical land area when initial corruption level and growth in real GDP per capita growth are partialled out.




Panel A: Geographical tropics

Panel B: Ecological tropics

Notes: e(DEGOV | X) is the residuals of an OLS regression of DEGOV on a constant, CCI in 1996 and GYCAP. e(TROPICS | X) is the residuals from a regression of the share of land area in the tropics on a constant, CCI in 1996 and GYCAP. The plot is then constructed as a scatter plot of the two vectors of residuals. Full lines are the associated simple regression lines.

Table 5 reports results from 2SLS. Several things should be noted. Firstly, all *DEGOV* point estimates in panel B are estimated with high precision. Save for column (5), where the $t$-value is roughly 1.9, coefficient estimates are more than 2 standard errors above zero. Secondly, all point estimates are quite stable regardless of which instrument is used.[26] Thirdly, coefficient estimates are in the range 0.040 to 0.064, which is two to three times the size of the corresponding OLS estimates. This may be an indication of measurement

[26] This is to be expected as the two instruments are highly correlated (correlation coefficient is 0.81).

error in the e-government variable with resulting downward bias in OLS results. As is well-known, under classical errors-in-variables OLS is biased towards zero (attenuation bias). For present purposes, OLS would then underestimate the causal effect. However, the Hausman specification test does not detect a statistically significant difference between OLS and 2SLS results, which suggests that endogeneity is not likely to be a pertinent issue. Finally, including or excluding growth in real GDP per capita, $GYCAP$, makes no statistically significant difference to the size of the $DEGOV$ slope estimate.

- Table 5 about here -

Even in large samples, IV methods can be ill-behaved when instruments are weak. Instruments with a low partial correlation with an endogenous variable can lead to severe bias in 2SLS. With one endogenous variable, the well-known "rule of thumb" states that we should not worry about weak instruments when the $F$ statistic from the first-stage regression is larger than 10 (Staiger and Stock 1997).[27] In table 5, instruments are strong in all columns.

If we take the most conservative 2SLS estimate, the economic significance of e-government is doubled. What does the 2SLS-based economic significance imply for Jamaica, the average country in the estimation sample. A one standard-deviation increase in $DEGOV$ increases the $CCI$ score with roughly 0.6 standard-deviation units, a bit more than twice the standardized coefficient in table 3, column (1). In the estimation sample, this would move Jamaica from $CCI$ country-ranking 76 to country-ranking 48 in 2006.

#### 4.2.1 Additional 2SLS issues

I have relied on just identification throughout. However, with over-identification one can effectively test whether instruments are correlated with the structural error. At first glance,

[27] The null under the $F$ test is that the instrument is zero in the first-stage regression. Specifically, the null is that the correlation between the instrument and e-government is zero once we partial out the effect of the other explanatory variables.

it would seem worthwhile to bring in additional instruments. For instance, one could rely on the *initial population density*. The basic idea is that in areas with a high population density, knowledge about new ICT technologies and how to use them will spread more quickly (Forman et al. 2005). Using the additional instrument in column (1), table 5, the model passes the OID test by a wide margin (OID $p$-value = 0.96) and 2SLS with two instruments yields almost exactly the same results as those reported in table 5 ($DEGOV$ coeff. est. = 0.43, robust std. err. = 0.014). However, little is probably gained by bringing in the additional instrument. First, the instrument is not *deeply* exogenous as opposed to land area in the tropics. Second, it is well known that OID tests have very low power. That is, the actual size of the OID test in small samples far exceeds the nominal size, implying that the test rejects too often (Hayashi 2000).

Finally, it should be note that the percentage land area in the tropics and the population share in the tropics are highy correlated (correlation coefficient = 0.98). Using one or the other makes no difference to the results reported in this paper.

# 5 Concluding Remarks

This paper makes two contributions. The first is to provide an opening attempt at systematically addressing the claim that e-government reduces corruption. The second is to propose a novel identification strategy based on a link between climate-related circumstances on the one hand and modern electronics and the quality of power supply on the other hand.

The paper documents a strong and robust positive partial correlation between increases in the use of e-government and decreases in levels of corruption over the 1996-2006 period. 2SLS results are shown to be consistent with OLS findings. In addition, standardized coefficients demonstrate that the size of the relationship is economically interesting. In sum, both the anecdotal evidence discussed in the Introduction and the empirical analysis provided in this paper support the view that e-government technologies are important tools in the struggle to reduce corruption.

## References


[1] Bardhan, P., 1997. Corruption and Development: A Review of Issues. *Journal of Economic Literature* 35(3): 1320-1346

[2] Chawla, R., Bhatnagar, S., 2004. Online Delivery of Land Titles to Rural Farmers in Karnataka, India. Available online at: <http://www.worldbank.org/wbi/reducingpoverty/case-India-BHOOMI.html>

[3] Chisholm, W., and Cummins, K., 2006. On the Use of LIS/OTD Flash Density in Electric Utility Reliability Analysis. Paper presented at *The Lightning Imaging Sensor International Workshop*. Available online at: <http://lis11.nsstc.uah.edu/Papers/Proceedings.html>

[4] Encyclopædia Britannica 2007. Thunderstorm, Article retrieved from *Encyclopædia Britannica Online*, September 23, 2007

[5] Forman, C., Goldfarb, A., and Greenstein, S., 2005. Geographic Location and the Diffusion of Internet Technology. *Electronic Commerce Research and Applications* 4(1): 1-13

[6] Hayashi, F., 2000. Econometrics. *Princeton University Press*

[7] IEEE 2005. *How to Protect Your House and Its Contents From Lightning*. IEEE Guide for Surge Protection of Equipment Connected to AC Power and Communication Circuits. IEEE Press

[8] Kaufmann, D., Kraay, A., Mastruzzi, M., 2007. Governance Matters IV: Governance Indicators for 1996-2006. Available online at: <http://www.worldbank.org/wbi/governance/wp-governance.html>

[9] Krasker, W., E. Kuh and R. Welsch, 1983. *Estimation for Dirty Data and Flawed Models*. Handbook of Econometrics. Volume 1, Chapter 11, Edited by Z. Griliches and M. D. Intriligator

[10] Kressel, H., 2007. *Competing for the Future: How Digital Innovations are Changing the World*. Cambridge University Press

[11] Lau, T., Aboulhoson, M., Lin, C., Atkin, D., 2007. Adoption of E-government in Three Latin American Countries. forthcoming in *Telecommunications Policy*

[12] Mauro, P. 1995. Corruption and Growth. *Quarterly Journal of Economics* 110, 681-712

[13] NASA 2007. Lightning Detection From Space: A Lightning Primer. Available online at: <http://thunder.msfc.nasa.gov/primer>

[14] Sachs, J., 2000. Tropical Underdevelopment. CID Working Paper No. 57, Center for International Development, Harvard University

[15] Staiger, D., Stock, J., 1997. Instrumental variables with Weak Instruments. *Econometrica* 65: 557-586

[16] Svensson, J., 2005. Eight Questions about Corruption. *Journal of Economic Perspectives* 19: 19-42

[17] WDI, 2004. World Development Indicators 2004. World Bank

[18] Wescott, C., 2003. E-government to combat corruption in the Asia Pacific Region. Available online at: <http://www.adb.org/Governance/egovernment_corruption.pdf>

[19] West, D., 2005. *Digital Government*. Princeton University Press

[20] West, D., 2006. Global E-Government, 2006. Available online at: <http://www.insidepolitics.org/egovt06int.pdf>

[21] Wooldridge, J. 2000. *Introductory Econometrics*. South Western

[22] World Bank, 2004. Making Services Work for the Poor. *World Development Report*

## Table 1: Summary statistics and correlation matrix

| Panel A: Summary statistics | | | | | |
|---|---|---|---|---|---|
| | Obs. | Mean | Std. Dev. | Min | Max |
| DCCI | 153 | -0.02 | 0.47 | -1.54 | 1.03 |
| DEGOV | 187 | 27.44 | 6.80 | 8.00 | 60.30 |
| CCI 1996 | 153 | -0.03 | 1.06 | -2.08 | 2.30 |
| GYCAP | 166 | 0.21 | 0.22 | -0.43 | 0.88 |
| GEOTROPICS | 150 | 0.49 | 0.48 | 0.00 | 1.00 |
| ECOTROPICS | 164 | 0.31 | 0.41 | 0.00 | 1.00 |
| **Panel C: Correlation matrix (obs=126)** | | | | | |
| DCCI | 1.00 | | | | |
| DEGOV | 0.08 | 1.00 | | | |
| CCI 1996 | -0.32 | 0.52 | 1.00 | | |
| GYCAP | 0.28 | 0.15 | -0.07 | 1.00 | |
| GEOTROPICS | -0.11 | -0.57 | -0.46 | -0.35 | 1.00 |
| ECOTROPICS | -0.10 | -0.43 | -0.38 | -0.25 | 0.81 |

## Table 2: Partial correlation between e-government and changes in corruption

| | | | |
|---|---|---|---|
| **Panel A: OLS** | | | |
| | ***Dependent variable: DCCI*** | | |
| | (1) | (2) | (3) |
| DEGOV | 0.021 | 0.020 | 0.024 |
| | (0.008) | (0.009) | (0.009) |
| CCI 1996 | -0.168 | -0.174 | -0.192 |
| | (0.042) | (0.045) | (0.045) |
| GYCAP | | 0.435 | |
| | | (0.200) | |
| R-squared | 0.11 | 0.20 | 0.16 |
| Total observations | 149 | 136 | 136 |
| **Panel B: LAD** | | | |
| | ***Dependent variable: DCCI*** | | |
| | (1) | (2) | (3) |
| DEGOV | 0.017 | 0.019 | 0.017 |
| | (0.008) | (0.010) | (0.009) |
| CCI 1996 | -0.104 | -0.125 | -0.113 |
| | (0.049) | (0.066) | (0.057) |
| GYCAP | | 0.079 | |
| | | (0.291) | |
| Pseudo R-squared | 0.04 | 0.05 | 0.05 |
| Total observations | 149 | 136 | 136 |

Notes: Dependent variable is the change in the Control of Corruption Index (CCI) from 1996-2006. All regressions include a constant. Standard errors reported in parenthesis. Panel A: Ordinary Least Squares (OLS) with robust standard errors. Panel B: Least Absolute Deviations (LAD) with bootstrapped standard errors (10,000 replications).

## Table 3: Standardized coefficients

| | (1) | (2) | (3) |
|---|---|---|---|
| DEGOV | 0.282 | 0.294 | 0.348 |
| CCI1996 | -0.377 | -0.411 | -0.452 |
| GYCAP | | 0.212 | |

Notes: Standardized coefficients associated with OLS regressions in table 2, panel A.

## Table 4: Robustness check of partial correlation between e-government and changes in corruption

**Panel A: OLS**

| | *Dependent variable: DCCI* | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| DEGOV | 0.015 | 0.021 | 0.019 | 0.028 | 0.023 | 0.021 | 0.017 |
| | (0.009) | (0.008) | (0.008) | (0.009) | (0.009) | (0.009) | (0.009) |
| CCI 1996 | -0.12 | -0.173 | -0.161 | -0.205 | -0.165 | -0.171 | -0.198 |
| | (0.044) | (0.043) | -0.045 | (0.043) | (0.042) | (0.045) | (0.061) |
| Excluded region | ssa | soa | mena | eap | na | lac | eca |
| R-squared | 0.06 | 0.12 | 0.11 | 0.17 | 0.12 | 0.12 | 0.10 |
| Total observations | 115 | 144 | 131 | 130 | 147 | 124 | 103 |

Notes: Dependent variable is the change in the Control of Corruption Index (CCI) from 1996-2006. All regressions include a constant. Robust standard errors reported in parenthesis. The regional key is as follows: East Asia & Pacific (eap), Europe & Central Asia (eca), Latin America & the Caribbean (lac), Middle East & North Africa (mena), South Asia (soa), Sub-Saharan Africa (ssa) and North America (na).

## Table 5: 2SLS estimates of e-government on changes in corruption

| | | | | | | |
|---|---|---|---|---|---|---|
| **Panel A: First stage regression** | | | | | | |
| | *Dependent variable: DEGOV* | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| CCI 1996 | 2.506 | 2.645 | 2.551 | 2.402 | 2.665 | 2.531 |
| | (0.509) | (0.571) | (0.542) | (0.408) | (0.456) | (0.445) |
| GYCAP | | 1.815 | | | 3.308 | |
| | | (1.950) | | | (1.892) | |
| GEOTROPICS | -5.001 | -4.613 | -4.981 | | | |
| | (1.039) | (1.264) | (1.123) | | | |
| ECOTROPICS | | | | -4.322 | -3.768 | -4.360 |
| | | | | (1.013) | (1.171) | (1.117) |
| F-test (strength of identification) | 23.2 | 13.3 | 19.7 | 18.2 | 10.4 | 15.2 |
| **Panel B: Second stage regression** | | | | | | |
| | *Dependent variable: DCCI* | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| DEGOV | 0.043 | 0.040 | 0.052 | 0.060 | 0.054 | 0.064 |
| | (0.016) | (0.017) | (0.017) | (0.023) | (0.029) | (0.024) |
| CCI 1996 | -0.264 | -0.263 | -0.312 | -0.295 | -0.292 | -0.328 |
| | (0.063) | (0.069) | (0.065) | (0.076) | (0.100) | (0.085) |
| GYCAP | | 0.314 | | | 0.249 | |
| | | (0.212) | | | (0.254) | |
| Hausman specification test (p-value) | 0.329 | 0.653 | | 0.161 | 0.603 | |
| Total observations | 138 | 130 | 130 | 142 | 130 | 130 |

Notes: Dependent variable in panel A is the change in e-government from 1996-2006. Dependent variable in panel B is the change in the Control of Corruption Index (CCI) from 1996-2006. All Two-Stage Least Squares (2SLS) regressions include a constant. Robust standard errors reported in parenthesis. The null under the Hausman specification test is that the difference in 2SLS and OLS coefficients is not systematic.