

Can We Measure the Power of the Grabbing Hand?

A Comparative Analysis of Different Indicators of Corruption

Alexander Hamilton

Craig Hammer



WORLD BANK GROUP

Development Economics
Development Data Group
January 2018

Abstract

Sustainable Development Goal 16 is explicitly committed to measuring aspects of corruption over time, and the identification of robust indicators to do so is an important endeavor. This paper critically reviews the strengths and weaknesses of various objective and subjective indicators of corruption, using the standard criteria of validity and reliability to identify indicators most salient to measuring Sustainable Development Goal 16. Consistent with the large literature in the field, the paper finds that the aggregate survey-based indicators of corruption, especially the Corruption Perceptions Index and the World Bank's Control of Corruption

indicator, despite some important reservations and limitations, are the most valid measures of the magnitude of overall corruption in many country contexts. However, in every case, the initial results using one indicator should be cross checked with the use of the other indicator, as there are some minor differences between how the two indicators are constructed, and in practice it is difficult to establish a priori which indicator is marginally more efficient. Furthermore, whenever possible, subjective indicators should be cross checked with objective indicators, even when the latter may be of a more narrow scope and time limited availability.

This paper is a product of the Development Data Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at chammer@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Can We Measure the Power of the Grabbing Hand? A Comparative Analysis of Different Indicators of Corruption ¹

Alexander Hamilton² with Craig Hammer³

JEL Codes: D72 D73 H11

Keywords: Governance Indicators; Anti-Corruption; Public Sector Governance; Public Sector Performance Measurement; Sustainable Development Goals

¹ This paper summarizes and adapts some of the key empirical findings emanating from Alexander Hamilton's doctoral thesis undertaken at the University of Oxford (2008-2012). The completion of this work would not have been possible without the professional and personal support received from the following: my supervisors, Raymond Duch, Indridi Indridason, and Desmond King; my examiners Pablo Beramendi, David Soskice, and Jane Green as well as the funding provided by the ESRC. Invaluable feedback was also provided by Cristina Corduneanu-Huci, Varvara Lalioti, the participants of Nuffield College's GRIPS Workshop and a seminar at the World Bank in Washington, DC.

² Department for International Development, 22 Whitehall, London, SW1A 2EG. UK Comments and feedback on this paper are welcome: alexander-hamilton@dfid.gov.uk.

³ The World Bank, 1818 H St NW, Washington, DC 20433, USA.

The role of corruption, the abuse of public office for private gain, in limiting human development is now well established. However, the ability to measure changes in corruption over time and across countries, in a valid and reliable manner, remains a contested issue. This is despite the proliferation of indicators purporting to do just that. The overarching concern of this paper is to critically evaluate the different measures or indicators of corruption that might logically be used to inform policy makers about changes in corruption across countries and over time. This evaluation is used in determining whether it is possible to associate changes in anti-corruption efforts/institutional features that may incentivize more or less corruption (independent variable of interest) to variations in reported or perceived corruption (dependent variable of interest).

The paper achieves this objective by critically reviewing the strengths and weaknesses of different objective and subjective indicators of corruption, in order to identify appropriate indicators. This objective is achieved by using the standard criteria of validity and reliability in order to discriminate between these potential indicators. This is an important exercise because the new Sustainable Development Goals are explicitly committed to measuring aspects of corruption over time, and therefore identifying robust indicators to do so will be critical to realizing this endeavor.

More specifically, this paper critically reviews the validity, consistency, and substantive focus of the following indicators of corruption: (1) the Corruption Perceptions Index (CPI); (2) the Control of Corruption Governance Metric (CC); (3) the Government Effectiveness Governance Metric (GE); (4) the Global Corruption Barometer; (5) the UN Survey of Crime Trends; and (6) evidence on the propensity of diplomats from different countries to break the law (hereafter ‘Tickets’). The paper finds, consistent with the large literature in the field, that the aggregate survey-based indicators of corruption, especially the CPI and the CC, despite some important reservations and limitations, are the most valid measures of the magnitude of overall corruption in many country contexts. In every case, though, initial results using one indicator should be cross checked with the use of the other indicator, as there are some minor differences between how the two indicators are constructed, and in practice it is difficult to establish, *a priori*, which indicator is marginally more efficient. Furthermore, despite a narrower focus, because comparative information on the tickets data is also available, it is also possible to cross check the results of these subjective

indicators with this objective (but more restricted) measure of objective cross-national corruption (albeit at only limited points in time).

The paper proceeds as follows: the next section, Section I, is a discussion of the main strengths and weaknesses, in terms of validity and reliability, of using the subjective-based versus the objective-based indicators of corruption to measure the dependent variable of interest (variation in actual corruption across time and countries). Section II is concerned with reviewing, in detail, the five major indicators of corruption, which have to a significant extent been found to be valid and reliable measures of this phenomenon. The paper concludes that the CPI and the CC are, broadly and with some caveats, good indicators for measuring variation in corruption in a valid and consistent manner.

I. Measuring Corruption: The Costs and Benefits of Using Survey-Based Indicators

In order to operationalize and use corruption indicators, it is necessary to identify a valid and reliable measure of corruption. That is, to identify a dependent variable of interest that has the following characteristics: (1) it is *substantively* focused on measuring corruption (*validity criterion*); and (2) it consistently measures this outcome across national contexts and/or across time (*consistency criterion*).

Broadly, attempts to measure corruption can be divided into two major categories: (1) objective indicators – which use ‘real’ data (costs, materials used, etc.) to calculate the magnitude of waste and abuse in public works and/or services; and (2) subjective indicators – which use survey data (of experts/elites/the public) to try and measure the perceptions and/or experience of corruption by different groups country specialists, business people, voters etc.).

Given that the focus of this paper is on trying to account for variation in corruption in general rather than in one specific policy domain (e.g. infrastructure or procurement), it will now be shown that the composite subjective measures of corruption are the most appropriate sub-set of indicators for this purpose. This is because, despite being based on perceptions, the major subjective indicators are highly correlated with narrower objective indicators, as well as outcomes associated with corruption (trust in government etc.), making it difficult to argue that objective indicators enjoy a ‘validity-based’ comparative advantage. Furthermore, subjective

indicators also enjoy an advantage, in that: (1) they are focused on overall levels of corruption, rather than a narrow subset of activities (e.g. corruption associated with bridge building etc.); and (2) they are more readily available and reliable with respect to measuring cross-sectional variation in (perceived) corruption, thus enabling cross-sectional regression analysis to be used. However, despite this comparative advantage of subjective indicators, objective indicators can still act as complementary robustness checks of any initial subjective-based results, thereby increasing confidence in these initial results as valid and reliable. By critically examining the comparative benefits and costs of subjective and objective indicators with respect to their (general) validity and reliability, it becomes possible to justify this strategy,⁴ as follows.

(A) The Costs and Benefits of Subjective Indicators

Validity (Benefits). Proponents of subjective indicators argue that despite methodological challenges, survey-based indicators can generate *valid* data regarding the magnitude of corruption. Via careful survey design and cross checks, it has been possible to ensure real perceptions are captured and the results of invalid survey instruments can be identified and eliminated via the construction of composite indicators (see next section). These arguments have been confirmed by empirical evidence. As such, indicators have been found to be robustly correlated with changes in outcomes theoretically associated with more/less actual corruption, such as economic growth (Mauro, 1995, 1998) or trust in government (Sandholtz and Koetze, 2000). These findings make it difficult to argue that just because subjective indicators do not measure revealed preferences, they are necessarily too flawed/noisy to be measuring what is actually occurring in practice.

In addition, subjective indicators have the advantage of capturing the complexity and interaction of different actors involved in corruption. Thus, the perceptions of policy makers, stakeholders and ordinary citizens can be documented, and the associations or lack thereof, between the responses of these actors can be used to test theories in a manner that narrower and objective indicators of outcomes cannot,

⁴ The issue of which subjective indicator is most substantively focused on corruption is addressed in the next section.

because they are focused on one policy domain/type of corruption.⁵ This advantage of survey based indicators in capturing the complexity of policy making has even been acknowledged by proponents of objective indicators (Golden and Picci, 2006) who recommend the continued use of subjective indicators when the research question is focused on broad or cross-sectoral corruption, something which cannot easily be gauged by objective indicators focused on narrower types of misappropriation.

Validity (Costs). Despite the ability to undertake careful survey design and cross checks, there is always the risk that subjective indicators can be problematic. Unlike objective indicators, they are not capturing the ‘costly to fake’ real actions of policy makers (e.g. different prices for medical supplies (Gray-Molina et al, 1999), cost/quality of road building (Olken, 2009), or the behavior of policy makers (Fishman and Miguel, 2007). That is, subjective indicators do not quantify what has actually occurred but rather what is perceived or assumed. This may be particularly the case if survey respondents have an incentive to strategically misrepresent their perceptions of corruption, or are simply not able to recall their experiences accurately due to perception biases (Bertrand and Mullainathan, 2001). However, as noted above, the fact that the composite indicators of corruption are highly correlated with objective outcomes theoretically associated with corruption, means that the magnitude of this problem should not be exaggerated. Empirically, as Kaufmann et al (2007) have shown, the major subjective composite indicators of corruption are highly correlated with objective measures, especially when the standard errors of both are taken into account. This makes it difficult to argue that the subjective indicators are seriously flawed to the point that they cannot predict the actual behavior of officials.

Reliability (Benefits). Proponents of subjective indicators argue that they can provide consistent cross-sectional information on corruption. By asking the same questions over time and across countries, to similar or the same target groups of respondents,

⁵ For example, when testing whether a more electorally accountable policy-making context reduces corruption, because of anticipated electoral reaction it is possible to use both elite perceptions – that may capture the magnitude of bribes key stakeholders have to make – as well as voter perceptions. This not only allows the testing of the equilibrium outcome – variation in elite bribes accounted for by the policy-making context – but also the micro-mechanisms of the theory. Do voters’ perceptions of the magnitude of corruption also vary by context? While it is theoretically possible for objective indicators to provide a measurement of actual corruption, they tend to be focused on one policy area and cross-national data and voter experiences tend to be limited/non-existent.

subjective indicators have ensured that there are reliable data sets of perceptions of corruption among different groups (voters, policy-making elites, etc.) in a large number of countries. This is obviously critical when undertaking comparative work that requires the existence of consistent measures across countries and/or time, in order to ascertain how variation in institutions, across countries, affects variation in corruption.

Reliability (Costs). Of course, survey-based indicators also face challenges. Composite indicators, such as the CPI, CC, and GE, require the use of multiple sources in order to identify and mitigate the effect of non-representative survey results. Unfortunately, this means that, over time, the indicators are unlikely to use the same sources to arrive at a given country's score. Thus, while using these indicators for cross sections is not a problem, their consistency over time is more questionable and such comparisons are generally discouraged (World Bank, 2012; Transparency International, 2011). However, the use of survey-based instruments for cross-sectional purposes has been shown to be robust, especially given how the large data sets allow for the inclusion of a large vector of control variables, and therefore have been extensively used in the empirical literature (see Persson and Tabellini, 2003).

(B) Objective Indicators

Validity (Benefits). The single biggest advantage of objective indicators of corruption is that they measure observed actions/outcomes and can thus quantify the abuse of public office for private gain. For example, Olken (2009), and Gray-Molina et al (2004) were able to quantify the extent to which resources devoted to road projects were actually diverted by public officials, and Fishman and Miguel (2007) were able to record the extent to which officials abused their diplomatic immunity. In short, the objective measures are able to provide a 'real number' that captures the magnitude of abuse of public office for private gain, in a way that subjective indicators may not.

Validity (Costs). Despite being able to come up with numerical values, the validity of objective indicators is not unproblematic. A particularly pertinent issue is that of external validity (Transparency International, 2011). Many of the objective indicators measure only one narrow type of activity (e.g. road construction or abuse of parking

tickets), which are not necessarily representative of across the board levels of corruption; that is the average propensity of policy makers to abuse their powers across policy fields. Even more problematically, objective indicators may not always be easy to interpret. For example, does a high bribery prosecution rate denote high corruption or a zero tolerance of corruption and an efficient criminal justice system (Lambsdorff, 2005)? Even when comparative objective data are available, then, the narrowness of the outcome measured and/or the difficulty of interpretation can make it difficult to argue that, in reality, the data are measuring the average propensity of corruption (Fishman and Miguel, 2006).

Reliability (Benefits). To the extent that objective indicators are measuring the same outcome across countries and/or time, they can be reliable. This is especially the case in natural experimental settings in which other factors are exogenous. Examining, then, how many tickets diplomats from different countries accrue in the same time period and legal framework can yield a reliable cross-sectional data set that isolates this behavior (propensity to abuse diplomatic privileges) while holding many other factors constant (same legal status and location).

Reliability (Costs). Objective indicators can be problematic if the contexts in which measurements are made are different and/or change over time. Comparing the cost of bridge building across countries, and then using this to infer the propensity for bribes and/or waste in other sectors may yield unreliable results if all other factors that determine costs are not identified and controlled for. This fact makes it problematic to develop a cross-sectional data set that captures variation in corruption in different policy-making contexts, or even in the same country over time. While very detailed objective indicators of corruption, with respect to large infrastructure projects, have been developed within specific national contexts (e.g. Olken, 2009, for Indonesia, and Golden and Picci, 2006, for Italy), developing comparative indicators for all countries can be unreliable because of the sensitivity of costs to a large number of other variables that cannot easily be identified and controlled for.

(C) Using Objective Indicators as a Complementary Robustness Check?

As the discussion of the comparative costs and benefits of subjective indicators vis-à-vis their objective counterparts suggests, the costs and benefits of each type of indicator may act as a check on the weaknesses of the other type of indicator. Specifically, subjective indicators may be able to measure broader types of corruption, but then it should be the case that, if they are valid, they will be highly correlated with narrower, objective indicators of corruption. Unsurprisingly, given this observation, and despite the continued and lively debate in the literature (see Tresiman, 2007, for a review), there is a growing consensus that both types of indicators can be valid, consistent, and can provide cross checks for each other. This is because a growing number of research projects have found that, ordinarily, objective and subjective indicators of corruption usually correlate quite strongly with each other, especially if the standard errors of both are taken into account (Kaufmann et al, 2007). Thus, initial skepticism towards subjective indicators is misplaced (e.g. Olken (2009)).

II. Operationalizing Corruption: The Different Indicators

From the universe of valid and consistent measures of corruption, it is now essential to critically review which one of these indicators is most likely to capture the type of ‘corruption that is of theoretical interest. Two indicators emerge as most likely to enjoy a comparative advantage in measuring broad corruption: the CPI and the CC.

(A) Measuring Corruption: Survey-Based Indicators

Since the mid-1990s there has been an exponential growth in the number of indicators attempting to measure corruption (see Treisman, 2007, for a review). Of particular importance has been the development of ‘composite survey based indicators’ of corruption, which try to utilize multiple (survey) sources to increase the accuracy of their measures. In accordance with most of the literature (e.g. Persson and Tabellini,

2003), three major composite indicators of corruption are used extensively in this analysis: (1) Transparency International's 'Corruption Perceptions Index' (CPI), which focuses on overall political corruption; (2) the World Bank's 'Control of Corruption' (CC) dimension of governance, which is a broader measure of public sector corruption; and (3) the World Bank's 'Government Effectiveness', a dimension of governance much more focused on the non-elected public sector.

(i) The Corruptions Perception Index

The first major, freely available, and cross-national measure of corruption is Transparency International's annual 'Corruption Perceptions Index'. The aim of this index is to measure the abuse of 'entrusted power for private gain' (Transparency International, 2011). Each country receives a score, which can range from 0 (extremely corrupt) to 10 (no corruption) and individual country scores are developed by aggregating and averaging normalized scores of 'corruption related data' emanating from a variety of sources.⁶ The aim of the index is to provide a measure of the extent to which public sector bureaucrats and politicians engage in corruption (Transparency International, 2011). It is thus a potentially noisy but valid measure of the level of corruption in a given country:

"The Corruption Perceptions Index...captures information about the administrative and political aspects of corruption."

Consistent with this definition, as Table 1 shows, most of the representative components of the CPI are concerned with capturing: both the abuse of public office by politicians (potentially elected) and unelected officials (e.g. 'bribing and corruption that exists in the public sphere'); see Table 1. Specifically, seven of the eight representative sources of the CPI are concerned with overall corruption (e.g. 'assessment of the pervasiveness of corruption among politicians and civil servants'; see Table 1), while one source is exclusively focused on the activity of politicians ('assessment of corruption in government'; see Table 1).

⁶ The 2010 scores were calculated using data from 13 different surveys or assessments produced by 10 independent organizations (Transparency International, 2011).

The CPI is available from 1995 when countries were first scored and has been updated yearly, so that in 2016, 176 countries and territories received a score. Over time, both the scope (number of countries) and the accuracy (standard error) of the index have improved. Furthermore, the CPI has been found to be highly correlated with measures of actual corruption (business regulation, public perceptions of corruption; see Treisman, 2007), thus suggesting that the CPI is effectively measuring an aspect of corruption.

Coding. As noted above, the CPI is an interval measure that ranges continuously from 0 (most corrupt) to 10 (least corrupt). The CPI is a composite index meaning that it utilizes a variety of different sources to arrive at each country's score (see Table 1). One of the essential criteria for a source to be used as part of the CPI is that it "...must provide a ranking of nations" (Transparency International, 2011). This criterion has the effect of precluding the use of sources that may provide scores for different countries, but do not use the same methodology (sampling frame etc.) across these countries. In short, all the sources used by CPI provide a consistent and comparative measure of perceived corruption⁷ at any given time. The way the composite index is developed is that each individual indicator of corruption is standardized (such that each source has the same weight) and then the average (mean) standardized score is calculated. Thus, the CPI score of a country in any one year is the (standardized) average score of all the sources available for that country.⁸ In order to reduce abrupt variations in scoring, the CPI actually tries to include sources from the last three years. This ensures that reliable scores for any given country are only developed if there are at least three sources available for that country in any given year.

Over Time Variation. Because the number of sources that can be used to construct the CPI can fluctuate over time (as sources must be current from the last three years and provide consistent comparative information), comparisons over long time periods are not advised, as the marginal change in the CPI over a given time period may reflect measurement error. However, because of the inclusion of sources from the last three

⁷ Transparency International also ensures that the definition of corruption used does not vary significantly by source (TI, 2000, p.6).

⁸ The standardization occurs in stages with each source.

years, it is possible to use the CPI year average over a short period of time (as in Persson and Tabellini, 2003) as a representative score for a country in a given time period.

Constituent Parts (data sources). As Table 1 indicates, the data sources of the CPI are numerous. The sources mainly consist of surveys of experts (smaller-N, e.g. the Economist Intelligence Unit) or surveys of business leaders (Institute for Management Development has a larger-N) and tend to be conducted by well-established international institutions, such as the Economist Intelligence Unit, the World Economic Forum, the World Bank etc. The advantage is that many of these groups (e.g. business leaders) are likely to experience actual high-level political corruption, although the CPI has the disadvantage of not ascertaining the perception of voters, except in a minor capacity.⁹

⁹ Since 2001 the CPI has generally not included surveys of the general public as part of its index.

Table 1: Representative Components of the CPI (2000)

Source	Who Was surveyed/asked?	Question/ Assessment (Bureaucratic/Both/Political)	Availability of data	Sample size (year)?
Political & Economic Risk Consultancy	Expatriate business executives	'Extent of corruption in a way that detracts from the business environment for Foreign companies'	1998, 1999, 2000 in 12-14 Asian countries	280 (1998) 700 (1999- app) 1027 (2000)
Institute for Management Development	Executives in top- and middle-management; domestic and international companies	'Bribing and corruption exists In the public sphere'	1998, 1999, 2000 In 46-47 countries	2515 (1998) 4314 (1999) 4160 (2000)
The Economist Intelligence Unit	Expert staff assessment	'Assessment of the pervasiveness of Corruption among politicians and civil servants'	2000 in 115 countries	NA (expert assessment)
International Crime Victim Survey	General public	'During 1999, has any government official In your own country, asked you to pay a bribe for his service?'	1999, 2000 in 11 countries	20,000 (1999) 20,000 (2000)
The World Bank & EBRD	Senior business-people	'State capture and frequency of irregular, additional payments to public officials'	1999 in 20 countries	3000 (1999)
Freedom House	US academics and Freedom House Staff	'Levels of corruption'	1998 in 28 countries	NA (expert assessment)
The World Economic Forum (Global Competitiveness Report)	Senior business leaders; domestic and international companies	'Irregular, additional payments connected with import and export permits, business licenses, exchange controls, tax assessments, police protection or loan application'	1998, 1999, 2000 in 53-59 countries	3167 (1998) 3934 (1999) 4022 (2000)

The World Economic Forum (African Competitiveness Report)	Senior business leaders; domestic and international companies	'How problematic is corruption? Irregular, additional Payments are required and large in amount'	1998, 2000 in 20-26 countries	582 (1998) 1800 (2000)
Political Risk Service	Expert staff assessment	'Assessment of "corruption in government"'	2000 in 140 countries	NA (expert Assessment)

Source: Transparency International, 2012

(ii) Government Effectiveness & Control of Corruption

Both of these measures are developed by the same institution (the World Bank) and rely on a common methodology. As such, it is possible to summarize the coding and some of the issues associated with each indicator jointly. It is also important to note that these two indicators form part of a larger set of 'Worldwide Governance Indicators' which also include: (1) Voice and Accountability, (2) Political Stability, (3) the Absence of Violence, (4) Regulatory Quality, and (5) the Rule of Law (World Bank, 2011). The World Bank uses 30 existing data sources to develop each of these indicators. The sources are selected to include the views of citizens, business owners, academics and experts drawn from the public, private, and NGO sectors from across the globe, and the standard methodology is used (World Bank 2011).

Coding. For all six indicators, the World Bank uses the same approach in order to develop an interval measure of the governance dimension of substantive interest. This entails standardizing the variables and then using an 'Unobserved Components Model' (UCM) to develop each indicator. This process therefore enables the development of the control of corruption and government effectiveness indicators that ranges from -2.5 (most corrupt/least effective) to 2.5 (least corrupt/most effective) (World Bank, 2011, online).

Over Time Variation. Due to the annual change in the number of sources over time, making inferences regarding the marginal change in a country's score over a short period of time is not advised (World Bank, 2011). Averaging the scores of countries over a few years to get a representative average for the time period is not problematic, due to the fact that sources from adjacent years are used to construct the indicator at any one time.

Constituent Parts (data sources). The two indicators are developed using a sub-component of the data sources, as they are available and applicable by year. The 30 data sources can be divided into: (1) surveys of households and firms (nine sources), (2) commercial intelligence information generators (e.g. the Economist Intelligence Unit, four sources); (3) NGOs (9 sources, including Freedom House) and public sector organizations (e.g. the World Bank).

(iii) Control of Corruption Cluster/Governance Dimension

The aim of this measure, like the CPI, is to capture the extent to which public policy makers abuse their public office for private gain. The aim of the index is thus (World Bank, 2010):

"...designed to capture...extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests."

Despite a formal definition that may appear to be more skewed towards political corruption, the CC, like the CPI, is primarily composed of sources, some of which are the same as the CPI, concerned with measuring both political and bureaucratic corruption. As Table 2 shows, three of the five representative sources used to construct the CC are concerned with general corruption (e.g. 'pervasiveness of corruption'; see Table 2); one source is focused on bureaucratic corruption ('an assessment of the intrusiveness of the country's bureaucracy'; see Table 2) and one source is concerned with political corruption ('is corruption in government widespread?'; see Table 2). In short, the CC is very similar to the CPI, although, unlike the CPI, one of its

representative sources is focused exclusively on corruption undertaken by bureaucrats, meaning that the skew towards unelected officials may be slightly greater vis-à-vis the CPI. Furthermore, the CC uses fewer representative sources (five versus eight), meaning that it may be slightly less likely to be composed of representative sources regarding the level of ‘political’ corruption vis-à-vis the CPI.

The index is available from 1996 and was published bi-annually until 2002, after which annual scores became available when countries were scored. Since then it has been updated yearly. Table 2 shows the component parts of the CC. Like the CPI, albeit to a lesser extent, the indicator is focused on surveys of experts, some of which are the same as the CPI, for example the Economist Intelligence Unit. However, it does, vis-à-vis the CPI, focus on both grand political and more petty corruption (e.g. level of petty, large-scale and political corruption).

Table 2: Representative Components of the CC

Source	Who Was surveyed/asked?	Question/Assessment of (Bureaucratic/Both/Political)	Source Type
Economist Intelligence Unit Risk-wire & Democracy Index	Expert Staff	‘Pervasiveness of Corruption’	Commercial Business Information Provider
World Economic Forum Global Competitiveness Report	Survey- Senior business leaders; domestic and international companies	‘Public trust in financial honesty of politicians Diversion of public funds due to corruption is common’ ‘Frequent for firms to make extra payments connected: (1) trade permits, (2) public utilities, (3) tax payments, (4) loan applications, (5) awarding of public contracts, (6) influence laws, policies regulations, decrees, (7) to get favourable judicial decisions.’	Non-Government Organization
Gallup World Poll	Survey- general public	‘Is corruption in government widespread?’	Commercial Business Information Provider
Institutional Profiles Database	Expert Staff	‘Level of petty, large-scale and political corruption’	Government
Global Insight Business Conditions and Risk Indicators	Expert Staff	‘An assessment of the intrusiveness of the country’s bureaucracy. The amount of red tape likely to countered is	Commercial Business Information Provider

		assessed, as is the likelihood of encountering corrupt of officials and other groups'	
--	--	---	--

Source: *The World Bank, 2012*

(iv) Government Effectiveness Cluster/Governance Dimension

Substantively, the GE captures the same underlying notion of the abuse of public power by policy makers as do the CPI and the CC. However, it is qualitatively different in that it measures the extent to which corruption takes place via *unit cost increases*, rather than the level of bribes. Thus, while measures that compose the CPI and CC are likely to capture the ability of public policy makers to rent-extract via extortion and bribes, the GE is more likely to capture the extent to which public officials abuse their office by reducing their workloads, gold plating their benefits, etc. The GE is developed using the same methodology as the CC but is qualitatively focused on bureaucratic corruption (World Bank, 2010, online):

“Government Effectiveness (GE) – [is designed to capture] perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.”

As Table 3 indicates, the GE is composed of almost the same representative sources as the CC. However, unlike the CC, the focus of the GE is very much on corruption in the bureaucracy (e.g. ‘bureaucratic quality’; see Table 3). In fact, only one of the five representative sources of the GE is not exclusively focused on bureaucratic corruption: ‘quality of the supply of public goods: education and basic health and the capacity of political authorities to implement reforms’ (see Table 3). Of course, if bureaucrats are accountable to elected officials, then it is anticipated that the GE will be highly correlated with both the CC and the CPI. While the GE, therefore, is substantively not focused on political corruption directly, it is likely to be correlated with measures of political corruption because of the relationship between politicians and bureaucrats.

Table 3: Representative Components of the GE

Source	Who Was surveyed/asked?	Question/Assessment of (Bureaucratic/Both/Political)	Source Type
Economist Intelligence Unit	Expert Staff	'Quality of bureaucracy / institutional effectiveness '	Commercial Business Information Provider
World Economic Forum Global Competitiveness Report	Survey- Senior business leaders; domestic and international companies	'Quality of general infrastructure Quality of public schools Time spent by senior management dealing with government officials '	Non-Government Organization
Gallup World Poll	Survey- general public	Satisfaction with public transportation system Satisfaction with roads and highways Satisfaction with education system	Commercial Business Information Provider
Institutional Profiles Database	Expert Staff	'Quality of the supply of public goods: education and basic health Capacity of political authorities to implement reforms'	Government
Political Risk Services International Country Risk Guide	Expert Staff	'Bureaucratic Quality '	Commercial Business Information Provider
Global Insight Business Conditions and Risk Indicators	Expert Staff	'An assessment of the quality of the country's bureaucracy. The better the bureaucracy the quicker decisions are made and the more easily foreign investors can go about their business. Policy consistency and forward planning How confident businesses can be of the continuity of economic policy stance - whether a change of government will entail major policy disruption, and whether the current government has pursued a coherent strategy. This factor also looks at the extent to which policy-making is far-sighted, or conversely aimed at short-term economic advantage.'	Commercial Business Information Provider

Source: The World Bank, 2012

(v) Other Subjective Indicators

As noted above, numerous subjective indicators of corruption are now being developed. However, focusing on the three aggregate indicators identified above is probably a more viable strategy than using any additional indicators. This is primarily because individual indicators may be more prone to being unrepresentative and/or measure corruption in a narrow manner. While the focus on composite indicators is justified in terms of validity and reliability, it is useful to briefly consider one narrow indicator, which may serve as a useful robustness check of the findings of any of the composite indicators.

The indicator in question is Transparency International's Global Corruption Barometer, which, since 2003, has accompanied the publication of the CPI. The logic behind the GCB is that it provides information on perceptions of corruption by the general public, whereas the CPI is almost exclusively focused on perceptions by elites (business people and experts).

Since 2003, Transparency International has, in collaboration with Gallup International, commissioned annual questions on different elements of corruption (Transparency International, 2011). The aim of the Global Corruption Barometer is to provide perceptions of corruption by the general public rather than by experts (the focus of the CPI). The questions are part of Gallup's Voice of the People survey (Transparency international, 2011). The number of countries the survey covers has varied over time (47-86), and in some countries the authorities have barred politically sensitive questions.¹⁰ As Table 4 shows, the sample survey is national in scope, although in some developing countries it is confined to major urban areas and conducted via mostly face-to-face or telephone interviews. The sample framework is either random or quota (varying by country) but is representative and large (40,838 respondents in 2003), although in every case the final results are weighted by demographic characteristics (age, groups, and sex) in order to make the results as representative of the general population as possible. The summary statistics of a typical GCB are shown in Table 4.

¹⁰ For example, in the 2004 survey only one of the 13 questions was allowed in the Arab Republic of Egypt.

Table 4: Descriptive Statistics of the Global Corruption Barometer (2003)

Number of Questions	Number of countries	Questions of Interest	Sample Size	Demographic Data controls
6	47	'Corruption is a significant problem in: political life' (no/yes, slightly/yes significantly)	40,838 (19,488) female (21,390) male	Age, Education attainment. Income level

Source: Transparency International, 2012

Substantively, what is interesting about this question is that it is narrowly focused on corruption at the highest policy-making levels, thus it is most likely to capture the perceptions of corruption of senior politicians and bureaucrats. Of course, given the fact that the survey data (1) are focused on only one type of respondent (the general public) who may not experience some types of corruption directly (e.g. demands for bribes in exchange for business licenses etc.); and (2) do not combine other indicators of corruption to eliminate/reduce the effect of unrepresentative results, means that the question would be problematic as the primary dependent variable.

Coding. Over time, the questions in the GCB have changed. For the purpose of this paper, the questions of interest were asked of respondents in the 2003 survey. Specifically, respondents were asked whether 'corruption had a not/significant/somewhat significant/very significant effect on (1) personal and family life, (2) the business environment, and (3) political life' (Transparency International, 2011). Each answer received a percentage of respondents, was interval in nature, and could vary from 0%-100%. There is significant variation in the perceptions of corruption across the different policy domains.

Over Time Variation. The questions asked have varied by year and in order to ensure the most valid questions are used, the data from the survey are based on the 2003 iteration. Due to the fact that the questions and sampling strategy change over time, it is not always possible to compare the results in this manner.

(B) Measuring Corruption: Objective Indicators

Objective indicators of corruption are less numerous and less standardized than the composite subjective indicators. Three types of objective indicators have generally been developed: (1) input-output analysis of the anticipated versus actual costs of construction/provision of services; (2) criminal statistics regarding the number of prosecutions for bribery; and (3) natural experimental data on the behavior of policy makers. As the research question of interest pertains to cross-national variation in the policy-making context, it is not possible to consider the efficacy of input-output analysis, as these data do not exist except for very detailed country case studies. However, comparative indicators for the other two objective types of data do exist and it is worth critically reviewing them.¹¹

(i) The United Nations Survey of Crime Trends and Operations of Criminal Justice Systems

This is a survey compiled and collected by the Crime Prevention and Criminal Justice Division of the United Nations. The survey began in 1970 and compiles annual data on the incidence of different types of crime in UN member states. The survey asks relevant public authorities in each member state to provide data, from their own national statistics, regarding the incidence of crime. Included in this survey are questions regarding the number of prosecutions for bribery per 100,000 of the population.

Coding. The rate of prosecution per 100,000 can, hypothetically, vary from 0 to 100,000 (interval range). As Table 5 indicates, there is significant variation in the number of prosecutions for bribery per 100,000 of the population for the year 2000, while on average there were 3.5 prosecutions for bribery per 100,000 of the population. The standard error is greater than the mean (9.6), the lowest per capita rate of prosecution is 0.01 per 100,000 (Pakistan), while the highest rate was in Romania (52.3 per 100,000).

¹¹ Another sub-set of objective indicators not examined are those based on the input-output analysis of major infrastructure projects. This is because such indicators have only been developed at the local/national level and cannot therefore be used to test theories linking changes in national level electoral contexts to the level of corruption.

Table 5: Prosecutions for Bribery (2000)

Number of Observations	Average number of prosecutions per 100,000	Standard Error	Minimum	Maximum
54	3.5	9.6	0.01	52.4

Source: United Nations, 2012

There are two major problems with using the bribery data as a measure of corruption. Firstly, there are little data from OECD countries and the definition of bribery varies by jurisdiction, so cross-country comparisons may be extremely problematic. Secondly, prosecution for bribery does not necessarily measure higher/lower levels of overall corruption. Countries may have low levels of prosecution due to the lack of corruption (e.g. Ireland), or due to the poor capacity of the legal system (e.g. Pakistan). Conversely, a high prosecution rate may indicate high levels of bribery (e.g. Romania) but also the use of prosecution as a deterrent (e.g. Hong Kong SAR, China). In short, because the use of prosecution may vary significantly by context (capacity, deterrent, etc.) it is not possible to use this indicator as a valid, cross-sectional measure of increased/decreased corruption (Lambsdorff, 2004).

Over Time Variation. Because of changes in the definition of bribery over time and the inconsistent use of the term across jurisdictions, both over time and cross-sectional analyses may be problematic. However, within-country comparisons may be possible: for example, among U.S. states in which issues of judicial capacity and the use of broadly similar definitions of bribery are likely.

(ii) Quasi-Experimental Data: The Behavior of Diplomats

While individual incidents of corruption may be opportunistic in nature, there is evidence that social, cultural, and institutional norms – which create incentives and expectations – may strongly condition the propensity for individuals to abuse public office for private gain, when the opportunity arises. Fishman and Miguel (2007) exploit the fact that UN diplomats in New York are immune from prosecution and use the

number of parking tickets issued to individual diplomats to develop a per capita measure of the abuse of parking violations. They divide the number of tickets issued to diplomats of a certain nationality between 1998 and 2000, by the number of diplomats in that country's UN delegation.

Coding. The number of parking tickets per capita (size of the diplomatic delegation) is an interval indicator ranging from 0-249. As Table 6 indicates, there is considerable variation in the number of tickets issued per capita, with the standard error (33.0) larger than the mean (19.7). For many countries, especially high-income OECD countries in Northern Europe, the number of parking tickets issued was 0, while as a region, the Middle East had the highest rate of ticketing (Kuwait had the highest rate of all countries: 249.4).

Table 6: Per Capita Issue of Tickets (1998-2001)

Number of Observations	Average per capita Issue of Tickets	Standard Error	Minimum	Maximum
137	19.7	33.0	0.0	249.4

Source: Fishman and Miguel, 2007

While this measure has several weaknesses – it may only be measuring a very narrow form of corruption norm, and diplomats are not necessarily representative of the population – it also has several strengths. Firstly, it does in fact correlate strongly with subjective survey data on corruption (Control of Corruption and the CPI), and secondly, diplomats may not be representative of the average citizen, but are more likely to be similar to the senior policy makers who may undertake large-scale corruption. The data are available for a large number of countries (146), comparable, and exist over a time period (1998-2001) for which corruption indices exists. Because of this, it is used as a robustness check for these subjective results.

(C) Which Measure?

Given that the aim of this paper is to examine variation in overall levels of corruption, the most valid and reliable measure of corruption is the one that is most likely to capture those elements of corruption associated with both politicians and the broader public sector.

From this, it is possible to argue that the CPI, CC, GCB, and Ticket data are most likely to satisfy these criteria, since these indicators appear to focus, to varying degrees, on general of corruption. Discriminating between these indicators is more difficult, because the way the indicators are constructed generates different costs and benefits. Narrowly focused indicators – such as the GCB and the Ticket data – are more likely to be less noisy indicators of one dimension of corruption. However, by focusing on a narrow range of actions or respondents, it may be the case that such indicators fail to capture the multi-dimensional nature of corruption. Conversely, the CPI and the CC have the advantage of combining multiple sources to provide a more comprehensive assessment of corruption, but may also be noisy as they rely on perceptions.

In fact, as the next section shows, using the CPI, CC, Tickets and GCB as dependent variables of interest does not alter the results, an unsurprising finding given the fact that these indicators are highly correlated (see next section). Despite this, it is possible to argue that the composite subjective indicators (the CPI and the CC) are the best starting point of any empirical analysis, as they are more comprehensive and thus more likely to capture all elements of corruption.

III. Do the Indicators Measure an Underlying Level of Corruption? An Empirical Assessment

The discussion above, as well as much of the literature regarding the validity of measuring corruption, suggests that both objective and subjective measures of corruption should be measuring the same underlying activity. The indicators are constructed using similar questions and the oversight capacity of politicians suggests that even indicators focused on different activities will be highly correlated. While correlation is not a measure of validity and cannot ascertain causality, a robust positive

association between the measures would increase confidence in the empirical strategy pursued; namely, to use different indicators of corruption to verify the initial results.

The strong correlation between the CPI and the CC has already been noted in several studies (e.g. Treisman, 2007), as has the strong correlation between the CPI and Tickets data (Fishman and Miguel, 2007). These studies have not only consistently found a robust association between these measures, but have also established that, as the number of sources used to create each indicator has increased over time, the correlations between the indicators have also become stronger. Once again, this does not in and of itself prove that the measures are valid, but it is consistent with the logic that as these measures become more efficient at measuring corruption, it would be expected that the correlation between them increases.

If representative measures of subjective corruption data are *consistently* measuring similar, but not identical, types of perceived corruption, it follows that such measures should be highly correlated. However, measures such as the CPI and the CC – which measure overall corruption– should be less strongly correlated with the GE, which is focused more on bureaucratic corruption. Furthermore, if the subjective measures of corruption are valid, we would also expect them to be negatively correlated with the objective measure of corruption. That is, less perceived corruption (higher CPI, CC and/or GE scores) should be negatively correlated with the per capita number of parking tickets issued to diplomatic teams at the UN.

Tables 7 and 8 show the correlations using the raw data (from the early to mid-2000s, so that all sources are from approximately the same time period) between all three subjective indicators¹² and the one representative objective indicator of corruption. As expected from the discussion above, each indicator is correlated with all other measures in the manner anticipated. Specifically, each of the subjective measures is highly correlated with the other two and it is not possible, even at the 1% interval, to reject the hypothesis that these indicators are measuring the same thing. Despite this extremely strong correlation, the CPI (more focused on broad corruption) is much more highly correlated with the CC (0.98) than with the GE (0.93) as would be expected, given their focus on slightly different aspects of corruption. In short, there is very little difference between the indicators, suggesting that they are reliable measures of

¹² Given the restricted sample size of the GCB, these data are not reported here so as not to significantly restrict the number of observations.

perceptions of corruption.¹³ Given that there is no agreed way to determine to what extent the substantive focus of the CC and the CPI exists, this is advantageous because it suggests that alternating between the two should not affect the results.

Furthermore, all three subjective indicators predict changes in the objective measure of corruption in the manner anticipated. Namely, higher scores on the subjective indicators are associated with a statistically significant lower incidence of issued parking tickets. This relationship is most robust when using the CPI (significant at the 1% interval) but also remains significant for the other two subjective measures of corruption (albeit at the 5% level). However, while the associations are robust, the magnitude of the relationship is not as strong (ranging from 0.19-0.30). This suggests that when using the pooled data, the relationship between the objective and subjective measures is noisy.

Focusing on the OECD sub-sample (Table 8) the results are largely similar. The subjective indicators are highly and significantly correlated with each other, with the only change being that the CC and the GE are now slightly more closely correlated (0.95 versus 0.93). Interestingly, the association between the subjective and the objective indicators is now considerably stronger. Specifically, the inverse relationship between the CPI and the per capita number of tickets issued is now significant at the 1% level, and twice as strong (0.60 versus 0.30). The relationships between the number of tickets issued and the other two subjective indicators is also stronger (0.58 versus 0.19 for the CC, and 0.62 versus 0.19 for the GE) and in the case of the CC, more robust (significant at the 1% confidence interval). Of course, it is not possible to deduce whether the imperfect correlation between the indicators is due to the fact that the tickets are measuring only a narrow type of corruption, or the subjective indicators are not capturing objective assessments. However, the fact that the associations are robust provides a basis for using these measures, even if they have to be subjected to exhaustive robustness tests.

Table 7 Correlation between Measures of Corruption (2000s data-Pooled)

Measure	CPI	CC	GE
CPI	-		
CC	0.98***	-	
GE	0.93***	0.93***	-

¹³ Although their validity cannot be verified by these correlations.

Tickets	-0.30***	-0.19**	-0.20**
----------------	----------	---------	---------

Note: CC and GE are inverted for clarity of interpretation. Based on 131 observations. Pair wise correlations. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level. Sidak test used to eliminate the possibility that robustness is due to multiple comparison fallacy. *Source:* Author's calculations using scores from 2010. When including the GCB data the sample is restricted to 27 observations. However, the indicator is highly correlated with the other measures.

Source: The Author

Table 8 Correlation between Measures of Corruption (2000s data- OECD Sub-Sample)

Measure	CPI	CC	GE
CPI	-		
CC	0.98***	-	
GE	0.93***	0.95***	-
Tickets	-0.60***	-0.58***	-0.62***

Note: CC and GE are inverted for clarity of interpretation. Based on 19 observations. Pair wise correlations. *** denotes significance at the 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level. Sidak test used to eliminate the possibility that robustness is due to multiple comparison fallacy. *Source:* Author's calculations using scores from 2010. When including the GCB data the sample is restricted to 27 observations. However, the indicator is highly correlated with the other measures.

Source: The Author

A second empirical test of the robustness of the different measures of corruption is the use of factor analysis. If the subjective and objective measures of corruption are measuring the same thing, we would expect them to load onto a latent (unobserved) measure of overall corruption. Specifically, we would expect the factor loadings of the subjective measures (given their scales, higher values denote less perceived corruption) to have factor loadings of the same sign and opposite to that of the objective measure (in which higher values indicate more corruption).

Table 9 shows that the factor analysis results yield consistent and significant values for both the pooled and the OECD sub-samples. Specifically, in both cases, the factor analysis yields only one significant (Eigen value >1.0) latent variable and in both cases the individual measures load onto this variable in the manner anticipated. Namely, the subjective indicators load onto the indicator positively, while the objective measure loads onto the latent variable negatively. This outcome is consistent with the expectation that higher scores on the subjective indicator and lower scores on the

objective indicator should load onto a latent variable measuring the overall good governance (absence of corruption) of a given country. The results are especially strong in the OECD sub-set with the objective indicator having a bigger, absolute loading (0.61 versus 0.30) onto the only significant latent variable.

Table 9 Factor Analysis of the Objective and Subjective Indicators of Corruption

Variable	Factor Loading for each measure (for 1 st Latent Variable)				Eigen value	Other Significant Latent Variables
	CPI	CC	GE	Tickets		
Pooled (n=131)	0.99	0.99	0.94	-0.30	2.95	No
OECD (n=19)	0.99	0.98	0.95	-0.61	3.25	No

Source: Author's calculations using scores from early 2000s.

IV. Conclusion

Can we measure corruption effectively? In order to test any hypothesis regarding the determinants of corruption and thereby monitor progress on how to reduce it, it is essential to ensure that an effective measure of the phenomenon can be identified and utilized. A good candidate indicator for this agenda must be substantively focused on the overall (vs. sectoral) level of corruption and be reliably measured over time. Using an extensive literature review, correlations, and factor analysis, this paper has shown that while both objective and subjective indicators of corruption meet these criteria, the most appropriate indicators are the composite subjective indicators: the CPI and the CC. This set of findings shows that it is possible to use robust subjective indicators of rent-extraction to measure underlying levels of corruption – an outcome that will be invaluable for measuring and monitoring progress against the Sustainable Development Goals.

Bibliography

Bertrand, M. And Mullainathan, S. (2001) "Do People Mean What They Say? Implications for Subjective Survey Data." *American Economic Review*, 91, pp. 67-72

Fishman, R. and Miguel, E. (2007). "Cultures of Corruption: Evidence From Diplomatic Parking Tickets." *Journal of Political Economy* 115(6), pp.1020-1048

Golden M, and Picci L. (2005). "Proposal for a New Measure of Corruption, Illustrated With Italian Data." *Economics and Politics*, 17, pp. 37-75.

Gray-Molina G., De Rada. E. P. and Yáñez E. (1999). "Transparency and Accountability in Bolivia: Does Voice Matter?" *Working Paper No. R-381*, Inter-American Development Bank, Washington, D.C.

Kaufmann, D. Kraay, A. and Mastruzzi, M. (2007). "The Worldwide Governance Indicators Project : Answering the Critics," *World Bank Working Paper: Policy Research Working Paper Series 4149*, The World Bank: Washington DC

Lambsdorff J. (2004). "Corruption: An Empirical Approach. " *American Journal of Economics and Sociology*, 61, pp. 829–853

Lambsdorff J. (2005). *Consequences and Causes of Corruption: What Do We Know from a Cross-Section of Countries?* Passau: University of Passau

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

Mauro, P (1998). "Corruption and the Composition of Government Expenditure," *Journal of Public Economics* 69(2), pp. 263-279

Olken, B. (2009). "Corruption Perceptions vs. Corruption Reality," *Journal of Public Economics*, 93, pp. 950-964

Persson, T. and Tabellini, G. (2003), *The Economic Effect of Constitutions* Cambridge (MA): MIT Press.

Sandholtz W and Koetzle W. (2000). "Accounting for Corruption: Economic Structure, Democracy, and Trade," *International Studies Quarterly*, 44, pp. 31-50

Transparency International (2011). "Corruption Perceptions Index; Global Corruption Barometer" Available at:

< <http://www.transparency.org/> >

< http://www.transparency.org/policy_research/surveys_indices/gcb >

Transparency International (2012). "What is the Corruption Perceptions Index?"

Available at:

< <http://www.transparency.org/> >

< https://www.transparency.org/cpi2012/in_detail >

Treisman, D. (2007). "What Have We Learned about the Causes of Corruption from Ten Years of Cross-National Empirical Research?" *Annual Review of Political Science*, 10, pp. 10-52

United Nations (2012). "The United Nations Survey of Crime and Operations of Criminal Justice Systems." Available at:

< <http://www.unodc.org/unodc/en/data-and-analysis/United-Nations-Surveys-on-Crime-Trends-and-the-Operations-of-Criminal-Justice-Systems.html> >

World Bank (2011). "Good Governance Indicators." Available at:

< <http://info.worldbank.org/governance/wgi/index.asp> >

World Bank (2010). "List of Developed Democracies and Why It Matters." Available at < <http://richleebruce.com/economics/1st-world.html> >

World Bank (2012). “World Development Indicators.” Available at:
< <http://data.worldbank.org/data-catalog/world-development-indicators> >