Government Transparency Institute

Mihály Fazekas [1], *Romain Ferrali* [2]

Advances in measuring corruption and agenda for the future

Working Paper series: GTI-WP/2023:01

March 2023, Budapest, Hungary

[1] Central European University and Government Transparency Institute
[2] Aix-Marseille School of Economics

Abstract

Measuring corruption is indispensable for identifying effective anticorruption policies and tracking progress towards lower levels of corruption. While there is a widespread perception that researchers and policymakers lack adequate quantitative corruption indicators, we argue that in fact there is an abundance of such metrics, although gaps remain. Tried-and-tested indicators range from expert ratings through national representative bribery surveys until transaction-based proxy indicators such as public procurement risk indicators. However, the diversity of measurement instruments and corrupt behaviours tracked creates new challenges: selecting indicators congruent with research and policy objectives, and combining different indicators into a coherent assessment (e.g. composite scores). This article reviews a series of state-of-the-art indicators and provides guidance to researchers and policymakers in selecting indicators appropriate to diverse use-cases.

1. INTRODUCTION

There has been a growing interest in measuring corruption both by researchers and policy makers in the last few decades. For example, Google Scholar returns over 30,000 entries for corruption and measurement keywords. Similarly, major policy actors have engaged with corruption, emphasizing the need for innovative measurement in just the last few years (e.g. European Commission Rule of Law reports[1], the International Monetary Fund's Anti-Corruption Challenge[2], the World Bank's renewed focus on anti-corruption[3], and the US government's push for better data-driven anti-corruption tools[4]). This widespread interest is understandable, as good measures of corruption are key to tracking progress in anti-corruption efforts, identifying causes of corruption, and hence developing more effective anti-corruption policies. Internationally comparable indicators have also great merit for drawing inferences from comparing countries over time.

Unfortunately, discussions around corruption measurement and the adequate uses of corruption indicators have been hampered by a series of popular (mis)conceptions: i) corruption cannot be measured; ii) corruption is/can only be measured using perceptions of corruption, as for example by Transparency International's Corruption Perceptions Index (CPI) or other widely used perception surveys; or iii) corruption can only be/is best measured at the country level, again as in the CPI. Perhaps furthering the confusion, popular, high-level conceptualizations of corruption, such as the "misuse of public office for private gain" (Rose-Ackerman, 2008) provide useful general guidance but are far removed from the specificities and diversity of actual corrupt transactions. Such broad-brush definitions coupled with blunt perceptions-based indicators allow for cross-country comparisons over time, but gloss over crucial differences among forms (e.g. grand vs petty corruption) and loci (e.g. sectoral or regional differences) of corruption (Heywood, 2017). Moreover, many conceptualisations and hence measurement of corruption mix the determinants of corruption with corruption itself; for instance, when an indicator measures transparency or conflict of interest rules instead of measuring actual corruption.

This article sets out to bust those myths about quantitative corruption measurement. Corruption, we argue, can be measured with acceptable accuracy, not only through perceptions, and not only at the country level. Our review of different corruption indicators reveals a dizzying number of measurement tools and concepts (for an overview see Table A1 in the Annex). While the availability of a rich set of tried-and-tested indicators is reassuring, such diversity might prove challenging. Hence, we offer guidance to researchers and policy makers for selecting indicators that are appropriate to their goals. This review article is not a classic literature review. Instead of trying to provide an exhaustive account of the literature, we offer a targeted focus on the main trends in the disciplines relevant for measurement by drawing on seminal articles and measurement tools while also highlighting some promising innovations. Hence, we selected indicators which are widely used and/or offer great promise

[1] https://ec.europa.eu/info/policies/justice-and-fundamental-rights/upholding-rule-law/rule-law/rule-law-mechanism/2021-rule-law-report/2021-rule-law-report-communication-and-country-chapters_en
[2] https://www.imf.org/en/Topics/innovation-at-the-imf/imf-acc-event
[3] https://documents.worldbank.org/en/publication/documents-reports/documentdetail/365421591933442799/anticorruption-initiatives-reaffirming-commitment-to-a-development-priority
[4] https://www.anticorruptiontechsprint.org/

and innovation, while being already systematically tested and repeatedly measured at least a few times.

This article covers the widely popular, older, country-level indicators based on expert ratings or nationally representative surveys, as well as more recent, more micro approaches that have recently gained in use. Expert ratings, such as the abovementioned CPI, are the first generation of corruption indicators. While their popularity has eroded over the last decade, an important reason for their success is their still unparalleled global coverage and the availability of long time series. There is also an increasing number of national representative surveys of citizens or enterprises which ask about corruption experiences and perceptions such as AfroBarometer. These surveys are typically run at regular intervals (sometimes annually), and allow for a granular assessment of bribery in different public institutions such as schools. The most recent innovations have focused on the micro-level, deriving new indicators measuring corruption at the transaction level (e.g., convictions) or at the individual actor level (e.g., high risk company ownership structure). Enforcement-based indicators have great promise in contexts where courts are independent and have reasonable capacity to go after corruption such as the US federal government. A wide class of indicators have also emerged that "approximate" corruption by trying to measure corruption risks instead of actual instances of wrongdoing. These track traces or red flags of corrupt transactions for example in public procurement or personnel moves (e.g. revolving door).

Our impressive arsenal for measuring corruption across the world, however, comes with a price: many indicators are costly to derive, posing issues of sustainability to some users. Furthermore, some gaps remain, with much room for improving and refining existing metrics. Finally, a few blind spots remain: systematic evidence about occurrences of corruption in public subsidies and grants, recurrent sale or renting of public assets, or welfare payments is seldom available.

Our guidance for users aiming to navigate this diverse landscape starts from the fact that indicators looking at the same type of corruption at the same level of observation tend to be highly correlated; however, they will considerably diverge as soon as the type of corruption or the level of observation changes. Given the many forms that corruption can take, such divergence is natural and, in fact, welcome. Researchers and policy makers need a diversity of measurement tools to serve their different needs rather than a single best indicator of corruption. The starting point for indicator choice is to define precisely i) the kind of corruption that is of interest and ii) the level of aggregation and degree of precision needed (e.g. macro versus micro levels).

The remainder of the article reviews each main indicator group, staring with the macro level - expert ratings and representative surveys - and continuing with micro level indicators such as law enforcement-based indices and proxy indicators. We conclude by offering guidance on indicator choice and proposing some fruitful avenues for further research based on identified evidence gaps.

2. THE MACRO-LEVEL: MEASURING CORRUPTION AT THE COUNTRY LEVEL

Macro-level indicators of corruption intend to summarize all dimensions of corruption occurring in a given location into a single measure. They represent the earliest attempts at providing quantitative estimates of corruption. They are ambitious, as corruption has many dimensions – for instance, bribery in the police is a different dimension than embezzlement in hospitals – that are each hard to measure.

Macro-level indicators are the most natural way to answer a series of questions, especially those that are inherently macro-level and cross-country. Such indicators are routinely used to provide descriptive evidence of the global incidence of corruption. For instance, Transparency International's (TI) Corruption Perception Index (CPI) is often used to issue global country rankings. Such indicators are also used to shed light on the macro-level causes and consequences of corruption. In a seminal paper, Mauro (1995) shows that high levels of corruption correlate with slower economic growth.

To answer such macro-level questions, one would need macro-level indicators that satisfy a series of criteria. Such indicators should (1) be comparable across time and space, (2) have broad coverage, both in terms of depth (time-coverage) and breadth (country coverage), and (3) measure corruption with no systematic bias and with as little noise as possible. This section assesses macro-level indicators of corruption according to these criteria. We will see that macro-level indicators tend to have the broadest coverage, both in terms of depth and breadth. They are, to some extent, comparable across time and space and, while assessing bias is difficult, they seem to be immune to some obvious sources of bias.

Macro-level indicators are grouped into i) expert ratings, derived using the judgement of a small number of experts, and ii) public opinion surveys, which gather corruption experiences or perceptions of a large number of citizens or companies. This section discusses each of these categories in turn.

2.1 Expert ratings

Expert ratings are, historically, the first indicators of corruption, emerging as early as the 1990s. Although there are many variants, these indicators are constructed by asking a series of experts their opinion on how corrupt a country is, perhaps along a series of dimensions, then aggregating across dimensions and across experts.

Three such indicators rank among the most popular and important: TI's CPI, the Varieties of Democracies (V-DEM) project, and the World Bank's Worldwide Governance Indicators (WGI). Alongside with measuring several dimensions of democracy, the V-DEM project also provides a political corruption index. This dataset uses at least 5 country-experts per country-year to measure corruption in the executive, legislative, judiciary, and public sector, and aggregates them into a political corruption index. The WGI provides measures of six dimensions of good governance, including "control of corruption," which will be our focus. The CPI and the WGI are not strictly expert ratings, in that they aggregate a series of other indicators (henceforth, "sources") into a composite indicator. While CPI uses expert ratings only, WGI also uses representative surveys.

The quality of these indicators largely depends upon whether experts' perceptions of corruption are accurate. Bias could emerge from some or all experts having inaccurate perceptions -- for instance, because they share a similar ideological background. While the former may introduce non-systematic bias to the data, the latter can lead to systematic bias. A natural way to assess whether these indicators are immune to bias would be to compare expert ratings to some objective measure of corruption, which proves hard to do at a cross-country, time-series scale (e.g. most countries in the world, for more than 25 years). It has been done in a within-country set-up, for well-specified types of corruption. For example, Banerjee&Pande (2007) finds strong correlation (0.75) between the expert-based and de facto measure of corruption among candidates for political office in Uttar Pradesh (India). However, such within-country and focused tests are unlikely to be representative of globally used expert assessments.

Systematic measurement bias can also be tested through measurement *reliability* and *credibility* (Rose&Peiffer 2019). An indicator is reliable to the extent that it correlates strongly with other tested indicators. An indicator is credible if it confirms existing intuitions about corruption. For instance, an indicator would not be credible if it implied that Norway is more corrupt than Nigeria. The CPI and WGI are very strongly correlated (0.98) and do not seem to have edge cases that would undermine their credibility. However, given the similarities among these indices, such high correlation is only a weak test of measurement reliability.

Since assessing the quality of macro-level corruption indicators remains difficult, attempts were made at addressing sources of bias individually. To minimize non-systematic bias, such datasets aggregate the ratings of several experts – typically (e.g., V-DEM), no less than five per country-year. Aggregation methods differ slightly, but all tend to move beyond simple averages to reduce the impact of an individual expert providing a dissonant rating.[5] As such, all these indicators come not only with a point estimate of corruption, but also with a confidence interval around it. Country-years for which fewer sources are available or sources disagree will have more uncertainty, reflected in wider confidence intervals, allowing researchers to discard observations with subpar measurement precision. In this regard, composite indicators (e.g., CPI and WGI) are particularly advantageous: pooling many sources increases the number of experts and the dataset's coverage.

The question remains whether these indicators display systematic bias stemming, for instance, from shared ideological views. Some early results are somewhat reassuring: evaluating the 2005 CPI ranking, Lambsdorff (2005, 2006) shows that Western experts' ratings correlate highly with non-Western ratings. In other words, if measurement is driven by Western understandings of corruption, these understandings must be so prevalent as to also inform the views of non-Western experts – a claim that is increasingly strong as the number of experts used to construct the dataset increases. However, experts might also share the same incorrect perception of corruption not because they hold the same worldview, but because they share the same sources of information. Olken&Pande (2012) illustrate this point using Indonesia's deteriorating CPI score (2.0 in 1998 to 1.7 in 1999) following Soeharto's death in 1998. They argue that it is not the actual increase in bribery which caused CPI to drop, but shifts in expert perceptions due to a freer press more extensively reporting on corruption. Although experts'

[5] V-DEM constructs 6 lower-level indicators using Bayesian Item Response Theory (IRT) models to aggregate experts' ratings, then takes a simple average of those lower-level indicators to construct the political corruption index. CPI (2021) uses 13 sources, aggregated using a simple average of the individual indicators, after standardizing them (Z-score transformation). WGI (2022) uses 23 sources, aggregated using an Unobserved Components Model (UCM) to construct a weighted average of the individual indicators.

ratings may indeed suffer from such biases, the rationale for relying on experts is precisely that they are more immune than the general public from such perception biases.

A final potential source of concern is that such ratings show little intertemporal variation. The 2015 and 2016 CPI rankings' correlation coefficient is 0.990, and the 2012 and 2016 rankings' correlation coefficient is 0.976. On the one hand, this may reflect bias from experts who update their perceptions slower than reality. On the other, overall corruption may well change slowly and substantial variation may take years. Hence, large score changes (e.g. Indonesia above) may reflect too strong or unfounded updating from experts.

Overall, expert ratings of corruption are presumably best suited to capture all dimensions of corruption in a single indicator. Their wide coverage makes them particularly well-suited for comparisons across countries, and investigating macro-level, long-run factors' impact on corruption (see Svensson, 2005; Treisman, 2007). Indeed, these indicators act under the premise that knowledgeable experts can report not only those small yet widespread instances of petty corruption, but also larger, less visible instances of grand corruption. As such, they are the only indicators that intend to cover all forms of corruption. Such a bold agenda inevitably comes with large measurement error, stemming from both systematic or non-systematic biases. A good practice is not to put too much faith in small differences, and systematically rely on the confidence intervals that come with those indicators to focus on the differences that are large enough to capture substantial variation in corruption.

2.2 Representative surveys

Representative surveys emerged after expert-based measures, in the early 2000s. Such surveys ask a large, representative sample of citizens or firms about their *experiences* of corruption, especially bribery[6], or about their *perceptions* of corruption; that is, how corrupt they think their country is. Measures of perceptions are typically more fine-grained, asking perceptions of corruption for a series of institutions, such as Parliament or police.

The Barometers family of surveys, TI's Global Corruption Barometer, and the WB Enterprise survey are prominent examples of such representative surveys; they ask about perceptions and/or experiences of corruption. The World Bank Enterprise surveys asks firm managers to estimate the value of bribes paid in "firms like yours" for a variety of transactions, such as paying taxes or securing government contracts. The family of Barometers surveys is a collection of loosely comparable surveys of citizens with a continental focus. It includes the AfroBarometer, ArabBarometer, AmericasBarometer, AsianBarometer and EuroBarometer. These surveys usually ask whether respondents have personally witnessed corruption or bribe-taking by politicians or government officials in the past year, and about respondents' perceptions of corruption in a variety of institutions. TI's Global Corruption Barometer asks similar questions to citizens and more rigorously comparable than the Barometers family. Finally, the European Quality of Government Index (Charron et al., 2014) measures both perceptions and experiences of corruption for selected public services (e.g. health, education) for EU countries, and is representative at the regional level, allowing for subnational comparisons.

Similar to expert ratings, one would like to evaluate the quality of these measures by comparing them to other indicators, including objective measures of corruption. (Treisman,

[6] Interestingly, crowd-sourced bribery information is increasingly available, e.g. I paid a Bribe in India: http://www.ipaidabribe.com/#gsc.tab=0 .

2007) compares expert ratings of corruption to citizens' perceptions and experiences of corruption. He finds moderate correlations ranging between 0.4 and 0.6 both within and between groups. Hence, according to Rose&Peiffer (2019), these measures are moderately reliable and credible.

Together, these findings show that expert ratings, perceptions, and experiences of corruption capture related, but different phenomena and should be evaluated differently. Experiences of corruption are objective, albeit local, measures of corruption, since they capture objective events occurring in one's live. Perceptions of corruption can be thought of as expert ratings carried out by presumably less informed "experts". However, such perceptions of corruption are interesting per se, as they inform other behaviors, in particular political behavior.

An important question with survey-based measures of corruption is whether respondents answer honestly these potentially sensitive question. The question is especially important for experiences of corruption, as misreported experiences is the most important threat to the measure. Some surveys try to mitigate the issue using impersonal wording; e.g., the WB Enterprise survey asks how much a "firm like yours" pays in bribe. A series of consistency checks suggest that reporting bias is negligible. (Rose&Peiffer, 2015 chapter 4) examine, among others, the AfroBarometer (2012 survey), AmericasBarometer (2012 survey), EuroBarometer (2013 survey), and the Global Corruption Barometer (2013 survey). They show that, in all surveys, non-response rates are low (< 10%), suggesting that few respondents find the question so sensitive that they would not answer. Furthermore, non-response rates do not correlate with respondents' moral stance towards corruption, nor with education.

An important question for perception-based measures of corruption is whether common citizens accurately perceive corruption. As with expert ratings, one would like to compare citizens' perceptions with objective measures of corruption. (Olken, 2009) conducts this exercise with road construction projects in Indonesian villages and finds that citizens perceptions of corruption are rather inaccurate. Olken's objective measure of corruption is the discrepancy between the costs reported by local governments and estimates of actual construction costs according to independent engineers. Citizen perceptions of corruption correlate with the objective measure, but the correlation is weak: a 10% increase of inflated costs on a road project increases the probability that a villager reports any corruption in said project by only 0.8%. Citizens' relatively less accurate perceptions of corruption may reflect different understandings of corruption. A large literature has examined the correlates and causes of perceptions of corruption. For instance, (Maeda&Ziegfeld, 2015) show that the level of education and income correlate with perceptions.

Taking stock, expert ratings, perceptions, and experiences of corruption capture different phenomena, making them more appropriate for some use-cases. Expert ratings are best suited to explore macro-level questions, as they aggregate all forms of corruption into a measure that has comparatively little bias, the widest coverage, and are most comparable spatially and temporally. Survey-based measures also have wide coverage and are relatively comparable spatially and temporally, although these advantages are less pronounced than for expert ratings: coverage is smaller and survey questions differ slightly across surveys. Moreover, these surveys allow for individual-level (citizen or firm) analysis of the causes and correlates of corruption.

Survey-based measures of corruption *experiences* seem to display little bias and therefore provide accurate measures of the *specific* forms of corruption respondents can reliably asess, such as bribery. Such indicators have often been used to analyse the micro-level determinants

underlying the incidence of corruption. In a seminal paper, Svensson (2003) analyses the incidence of bribery among Ugandan firms and shows that firms with low ability to pay and high “refusal power” are less likely to pay bribes.

Survey-based measures of corruption *perceptions* are more noisy than expert ratings. However, perceptions of corruption are an important driver of behaviour and have therefore generated much scholarly interest. A large literature has investigated the causes and consequences of perceptions of corruption, showing that poor perceptions of corruption undermine trust in government (e.g. Rose-Ackerman&Palifka, 2016; Rose&Peiffer, 2019), and their effect on the accountability of public officials (De Vries&Solaz, 2017).

3. THE MICRO-LEVEL: MEASURING CORRUPTION WHERE IT OCCURS

Micro-level indicators of corruption measure corruption at a small scale: either the agent (individual, firm, bureaucrat, politician) engaged in corruption, or the transaction during which corruption occurs. Those indicators typically rely on administrative data. As such, and contrary to macro-level indicators, which rely on the perceptions of experts or citizens, these indicators take a more objective lens on corruption. Furthermore, administrative data offers wide, consistent coverage, and high granularity. Finally, while macro-level indicators lump together many forms of corruption, micro-level indicators tend to measure one specific facet of corruption (e.g., corruption cases that were prosecuted, or instances of corruption occurring in public procurement).

Practitioners will typically rely on these indicators when interested in the specific facet picked up by an indicator. Since administrative data tends to be specific to the laws and regulations of the country that produces them, such indicators also tend to be best suited for within-country comparisons. Finally, those indicators are as good as the administrative data they rely on, prompting careful examination of potential biases. For example, not all corruption cases are prosecuted, and bureaucrats may have an incentive to misreport the details of some procurement contracts. In what follow, we review each family of indicators in turn, and highlight, when relevant, specific use-cases.

3.1 Measuring proven cases: Law enforcement-based indicators

The common trait of corruption indicators based on law enforcement and related data is that they all claim that corruption is identified by a (more or less) reliable body tasked with monitoring corruption following a clear definition and expectations of evidential proofs (see Kaufmann, Kraay,&Mastruzzi, 2006; Transparency International, 2012). Classics in this field use corruption convictions by courts, although audits and investigation results, including police or investigative journalists, are also widely used.

Corruption indicators using convictions by courts have been developed across the globe, both in national and international contexts (e.g. Lopez-Iturriaga&Sanz, 2018; Campos et al., 2019). One notable example is federal courts' prosecutions and convictions of federal employees in the US, with time series going back to 1986 (Cordis&Milyo, 2016). Such data is obtained from the federal government by the Transactional Records Access Clearinghouse (TRAC), using freedom of information requests. It contains information on key aspects of federal corruption cases such as convictions, prosecutions, prison sentence and its length, length of proceedings (which points at some of the time gap between the corrupt act and conviction), and level of the officials convicted or prosecuted. Interestingly, this hard measure of corruption correlates weakly with survey estimates of corruption. They also predominantly point at low-level, bureaucratic corruption rather than high-level political corruption. This may indicate that such an indicator is affected by selection bias in favour of cases which can pass the high burden of proof in federal criminal courts, given prosecutors' capacity constraints (i.e. low value corruption, which is easier to prove than high value corruption). There were also attempts to create cross-country corruption indicators using law-enforcement data. Most notably

Escresa&Picci (2015) use domestic courts' convictions of foreign bribery following the OECD Anti-Bribery Convention. While this indicator allows for cross-country comparisons, it suffers from a few shortcomings, including relatively small sample size and potential inconsistencies, as national courts may use somewhat different definitions of corruption.

Similar to convictions and prosecutions-based measures, a series of corruption indices use data from anti-corruption audits (Trapnell, 2015). Such measures are particularly well-developed for Brazil, where audits such as federal audits of municipal spending of federal funds are randomly conducted (Ferraz&Finan, 2008; Vaz Mondo, 2016). They have also been developed for other countries (Gerardino et al., 2017). While audits often do not identify corruption directly, they tend to unearth false payments, fraudulent claims, procedural irregularities, and unreasonable spending which are typically symptomatic of corruption, even if they may also result from incompetence.

Using suspected, albeit not proven cases of corruption, as reported by investigative media or NGOs is increasingly used for measuring corruption. Even if such suspicions are based on well-documented, hard facts, they may be less reliable than criminal convictions or audit reports. Such reports of corruption tend to be very rich in detail enabling not only identifying the extent of corruption but its structure, forms, and actors (Jancsics&Jávor, 2012). However, measurement might be biased by media's relationship with political actors, the public's interests (e.g. in sports), and quality of law enforcement leading to different representations of corruption even in similarly corrupt countries (Mancini et al., 2016). While some have attempted to gauge the quantity of corruption using suspicion-based data (e.g. Ferwerda et al., 2017), generally they are better suited for advancing our qualitative understanding of corruption.

Overall, enforcement-based indicators tend to be reliable for establishing and explaining petty and grand corruption as they derive from precise, legal definitions of corruption and measurement is usually done by organisations dedicated to finding corruption under public scrutiny. In other words, such indicators typically have few false positives; that is, we can be confident that instances where they detect corruption signal actual corruption (albeit suspicion-based indicators are likely to have higher false positive rates). In addition, the stability of law-enforcement bodies and definitions over time largely warrant building long time-series of micro-data such as in the US federal corruption convictions example.

However, these indicators are not without downsides. Sample selection biases arise in two forms: first, what gets audited, investigated or prosecuted is rarely random requiring an initial tip and often political support in high-level cases (e.g. courts sentencing high profile politicians for corruption just after they lost office may reflect power games rather than objective sampling of corrupt cases). Brazil's *random* audits are the exception rather than the rule in this field. Second, what can be proven as corruption may instil further bias, as it is tends to be narrow and depend on how sophisticated corrupt actors are. Outright corruption may also be hard to prove. Prosecution might attempt to prove other offenses instead (e.g., tax evasion), further biasing corruption measurement.

In addition to sample selection biases, the narrowness of legal definitions of corruption leads to a considerable mismatch with the corruption concepts relevant for theory or policy making. Furthermore, the often-formalistic nature of audits mean that only very apparent forms of corruption are identified in narrow administrative contexts. For example, the widely cited random audits of Brazilian municipalities only concern federal funds spent by local entities, hence represent a biased assessment of municipal corruption when municipalities can shift rent seeking from federal to local funds.

Taken together, the clear corruption definitions and low false positive rate of enforcement-based indicators make them attractive for corruption research. However, sample selection bias and narrow focus raise questions regarding their use for corruption research and policy analysis. These make them more useful for looking into the kinds of corruption that can be picked up by law enforcement, or analysing anti-corruption enforcement itself.

3.2 Corruption in public employment: favouritism and nepotism

In the last decade a research agenda has developed measures of corruption in public employment, capturing phenomena such as nepotism and favouritism. These indicators are measured on the level of individual bureaucrats, but are often aggregated to the organisational or regional levels. Good measurements exist for diverse forms of corruption in public employment using both administrative and survey data. For measurement, the starting point is meritocracy in appointment and promotion, rather than connections or bribes and independence of bureaucrats from political interference (Dahlström et al., 2012).

Indicators using administrative data capture the politicization of the civil service through political appointments to key civil service positions (Gordon, 2011; Hollibaugh et al., 2014). Such data usually does not allow for a direct measurement of corruption, providing instead indirect evidence of corrupt motives. For example, Fazekas et al (2022) show that presidential appointments in agencies weakly insulated from the president facilitate awarding government contracts to companies donating to the president's electoral campaign. Moreover, nepotism in public employment as direct evidence for corruption is increasingly measured recently. For example, measurement based on surnames of civil servants in Sweden show how the public administration became more autonomous and professional over the centuries (Sundell, 2014).

Other indicators come from civil service surveys, which provide a flexible measurement instrument. They have been widely used for over a decade now (Meyer-Sahling, 2009). Some surveys use direct questions of perceptions and experiences of corruption in the civil service, in a very similar fashion to representative surveys of corruption discussed above. They tend to suffer from the same kinds of reporting biases, although attempts have been made to alleviate such biases using list experiments and other indirect methods for eliciting sensitive answers (e.g. Schuster et al., 2020).

Overall, these indicators of corruption are well suited to study one specific form of corruption: favouritism in public employment. The kind of administrative data these indicators usually rely on offers wide, consistent coverage. However, such data usually offers indirect evidence of corruption, hence introducing some degree of imprecision, and may suffer from systematic biases, for instance in the way corruption is reported. Survey measures also suffer from the biases enumerated above (section 2.2).

3.3 Proxy indicators: tracking risks in micro-level transactions

Micro-level transactions-based indicators measure corruption on the level of economic transactions where corruption actually takes place. These indicators invariably resort to proxying corruption instead of measuring it directly, implying that establishing their validity relies on theory and statistical evidence. As corruption manifests itself in a range of economic

and political transactions, these indicators vary by type of transaction (e.g. welfare payments, government contracts), or by type of actors (e.g. companies and personal connections). We review them in turn.

3.3.1 Gap analysis of public spending

One increasingly popular proxy indicator approach is to identify spending gaps using different measurements aiming to capture the same or very similar sets of transactions (Sequeira, 2012). Such approaches either collate two administrative datasets (e.g. Klasnja, 2016) or an administrative dataset with an independent survey or external assessment (e.g. Reinikka&Svensson, 2004). Golden&Picci 's (2005) measure exploits differences between the quantity of infrastructure (stocks) and cumulative public spending on it (flows). Using Italian data, the authors provide suggestive evidence for the validity of their corruption measure by correlating it with Putnam's institutional performance scores for Italian regions (linear correlation coefficients range between 0.8 and 0.9 for various specifications and index variants). Unfortunately, Golden&Picci's indicator cannot differentiate between different types of corruption, nor separate corruption from wastage or inefficiency. Comparing official administrative records (e.g., welfare and within-government transfers) with end-user surveys is another route that has proven especially useful in developing countries, where administrative data tends to be weaker. Reinikka&Svensson (2004) compares Ugandan central government allocations to each primary school with the authors' own expenditure survey of the schools. Shockingly, they find that on average, schools received merely 13% of their allocations, with most schools receiving nothing. Olken (2006) uses a similar logic to proxy corruption when he compares government administrative records on subsidized rice allocations with a household survey data on actual rice received. A large portion of the allocations were actually received: on average 82%.

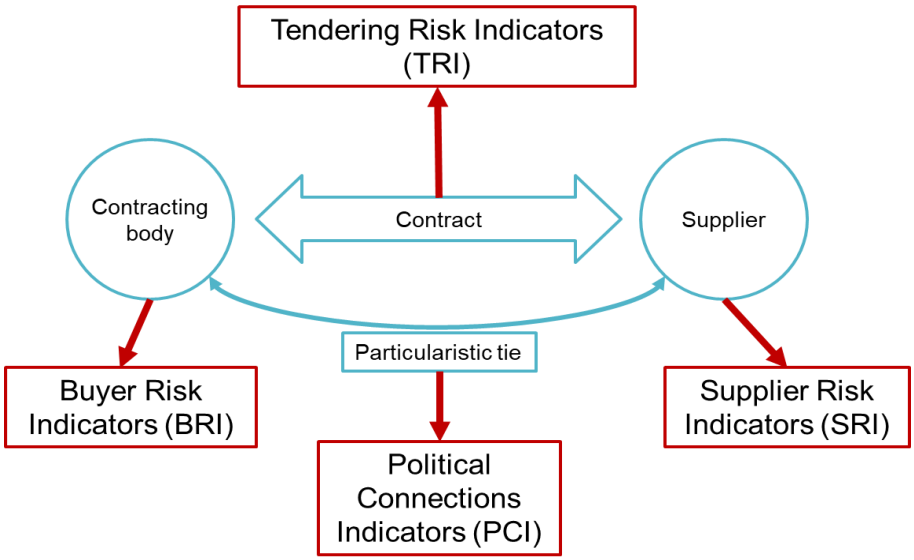
Common shortcoming of gap analyses is that they are unable to separate corruption from mismanagement (e.g. rice simply rotting rather than stolen). Moreover, identified spending gaps are likely due to a mixture of low- (e.g., individual bureaucrat administering the program stealing a small amount) and high-level corruption (e.g., a politician directing the program systematically shaving off a given percentage of all transfers). Overall, these indicators are well fitted to study the impact of corruption on overall state capacity, since they allow capturing the total extent of overspending in the implementation of a given policy.

3.3.2 Public procurement

A large body of research uses government contracting or public procurement data to develop proxies for corruption. The relative attractiveness of such approaches lies in the fact that public procurement data and regulations are relatively standard across many countries such as among European Union member states or the signatories of the World Trade Organization Agreement on Government Procurement. In addition, public procurement regulations around the world, by and large, aim to establish transparent and efficient markets which typically results in a lot of publicly available data. Hence, public procurement is a government function where valid corruption proxies can be constructed not only matching specific micro-contexts but also across many countries (Fazekas&Kocsis, 2020).

In public procurement, corruption proxies are very many and diverse, hence they are often grouped in sub-categories, for example according to the participating actors and their relationships (Fazekas et al., 2018): 1) tendering risk indicators characterising the contract award and implementation process, 2) supplier risk indicators describing the winning organisation, 3) buyer risk indicators characterising the buying organisation, and 4) political connections indicators capturing the links bridging the public-private divide (Figure 1).

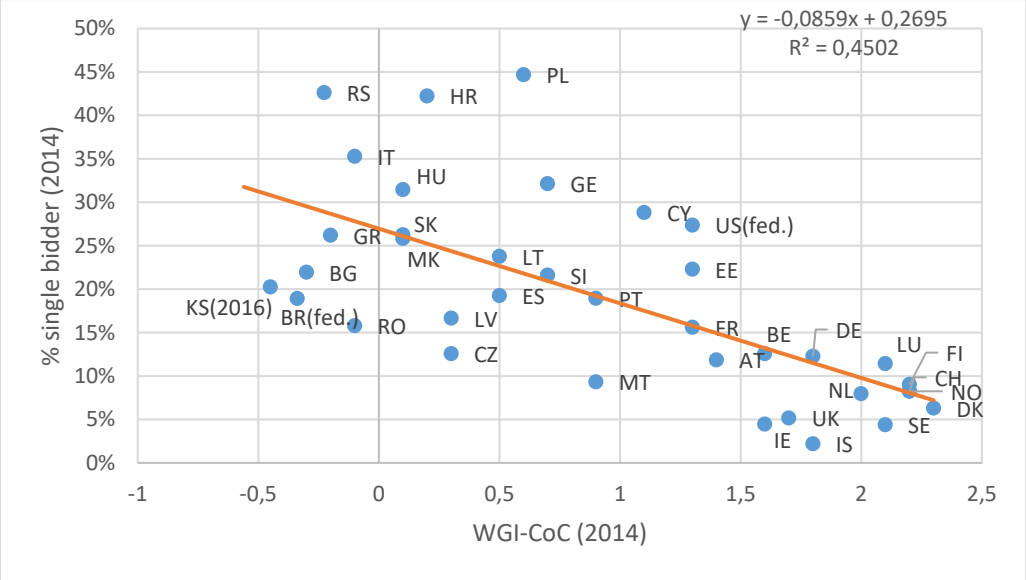
Figure 1. The 4 key elements of corrupt exchanges in public procurement and the corresponding indicator groups



Source : Adapted from (Fazekas et al., 2018)

Tendering Risk Indicators aim to capture the corrupt manipulation of the tendering process - starting from launching the tender to completing the contract implementation – serving the corrupt in generating rents and allocating them to the favoured companies (Fazekas&Kocsis, 2020). One notable example is the tailoring of tender specifications to fit the favoured company with the effect of permitting only one bidder to 'compete' while excluding other bidders who would have had the capacity to compete for the contract (Transparency International, 2006). This phenomenon is best captured by single bidding in competitive markets (i.e. only 1 bid submitted when there are known competitors active on the market) (Charron et al., 2017). It has a strong content validity, that is it has a close association with the definition of corruption as limited access to public resources (North et al., 2009) based on particularistic connections (Mungiu-Pippidi&Fazekas, 2020). Furthermore, there is substantial evidence for its convergent validity, i.e. its association with other indicators of corruption. First, other risk indicators characterising the tendering process, hence signal the deliberate manipulation of tenders to turn competition off for favoured bidders, also correlate with single bidding on the tender-level. Some of these indicators include the use of exceptional procedures (Auriol et al., 2011; Chong et al., 2015), lack of publication of the call for tenders (Fazekas et al., 2016), tailored participation pre-conditions (Decarolis&Giorgiantonio, 2019), and cost overruns (Olken, 2007). Second, corruption perceptions correlate with single bidding as well as its associated tendering risk indicators (e.g. lack of publication) on the country-year level (Figure 2). Third, supplier risk indicators such as tax haven registration also correlate with single bidding ad other tendering risk indicators. Fourth, overpricing is also strongly associated with single bidding and other tendering risk indicators. For example, single bidding is on average 10% more expensive than multiple bids in a Europe-wide sample of over half a million contracts in 2009-2014 (Fazekas&Kocsis, 2020).

Figure 2. Single bidding on competitive procurement markets and corruption perceptions (WGI-Control of Corruption index), 2014



Source: (Fazekas&Kocsis, 2020) *and authors' further compilation based on Government Transparency Institute data*

Supplier Risk Indicators signal the use of winner companies as vehicles of rent extraction facilitating the distribution and hiding of assets which are indispensable for rewarding the participants of the corrupt deal and avoiding detection (Fazekas&Tóth, 2017). They are also often related to company risk indicators discussed below. Probably one of the most widely quoted examples of this indicator group is registering the main supplier or some of its owners in a tax haven or secrecy jurisdiction (Christensen, 2011). Another widely used and intuitive indicator tracks company success when governments change aiming to gauge partisan favouritism benefitting companies with ties to the government of the day (David-Barrett&Fazekas, 2019).

Contracting Body Risk Indicators capture the weaknesses of administrative structures designed to insulate the whole organisation but most importantly the contracting function within from pressure to favour a bidder (Dahlström et al., 2021). Crucially, without the cooperation of bureaucrats administering tender and implementing contracts, corruption in procurement is unlikely to succeed. Most examples from this group build on concepts such as meritocracy and political insulation discussed in section 3.2, while applying them to the more specific context of public procurement.

Political Connections Indicators aim at tracking personal connections between the winning firm's owners or managers and key public officials able to influence the public procurement process. Connections can be based on a range of phenomena such as kinship, friendship, professional association membership, lobbying, or party donations (Fazekas et al., 2022). These indicators build on personal connections indicators discussed below, while applying them to the context of procurement.

While any of these individual indicators are likely biased, they can be aggregated to increase measurement validity (Fazekas&Kocsis, 2020). In our case, bias resides in the overreporting of false positives; that is, signalling corruption when the reasons are non-corrupt such as low

administrative capacity or exceptional circumstances. When indicators capturing corruption different phases of the procurement process (e.g. tendering outcomes and processes) and indicators from different indicator groups (e.g. tendering and supplier risks) are combined they can effectively lower false positive rates (Fazekas&Kocsis, 2020).

The diversity and international comparability of corruption proxies or red flags of corruption in public procurement make them one of the most widely used recent indicators. They are useful for studying micro-level relationships on the contract or organisation levels, but they are also amenable to cross-country comparative analysis of specific corrupt transactions. While there is good statistical evidence of their validity, individually and jointly, they suffer from imprecisions and biases. Most notably, it is often not possible to isolate alternative reasons for risky patterns, e.g. monopolistic markets or bureaucratic incompetence. With many innovations happening in this field (e.g. Decarolis&Giorgiantonio, 2019; Wachs et al., 2021), measurement accuracy is likely to improve.

3.3.3 Company risk indicators

Corruption proxies targeting companies have been developed touching on related topics such as financial secrecy, money laundering, and tax evasion while they have also been linked to public procurement, i.e. supplier risk indicators (see above). While the abuse of companies for corrupt purposes has been long recognised, we lack systematic evidence on which corporate characteristics signal corruption in which context. We highlight 3 indicator groups following the type of data used for indicator building: i) company registry information (e.g., many companies registered on the same address); ii) financial data (e.g., extreme profitability); and iii) ownership and management structure (e.g., hidden owners).

Registry information reveals fundamental attributes of companies, including headquarters location, size, and incorporation date. These attributes can be used to assess whether the company us anomalous compared to ‘clean’ businesses in a suitable benchmark market, hence proxy corruption. For example, case studies show that many companies involved in corruption are registered at an address where a great number of other companies are also registered (Caneppele et al., 2009).

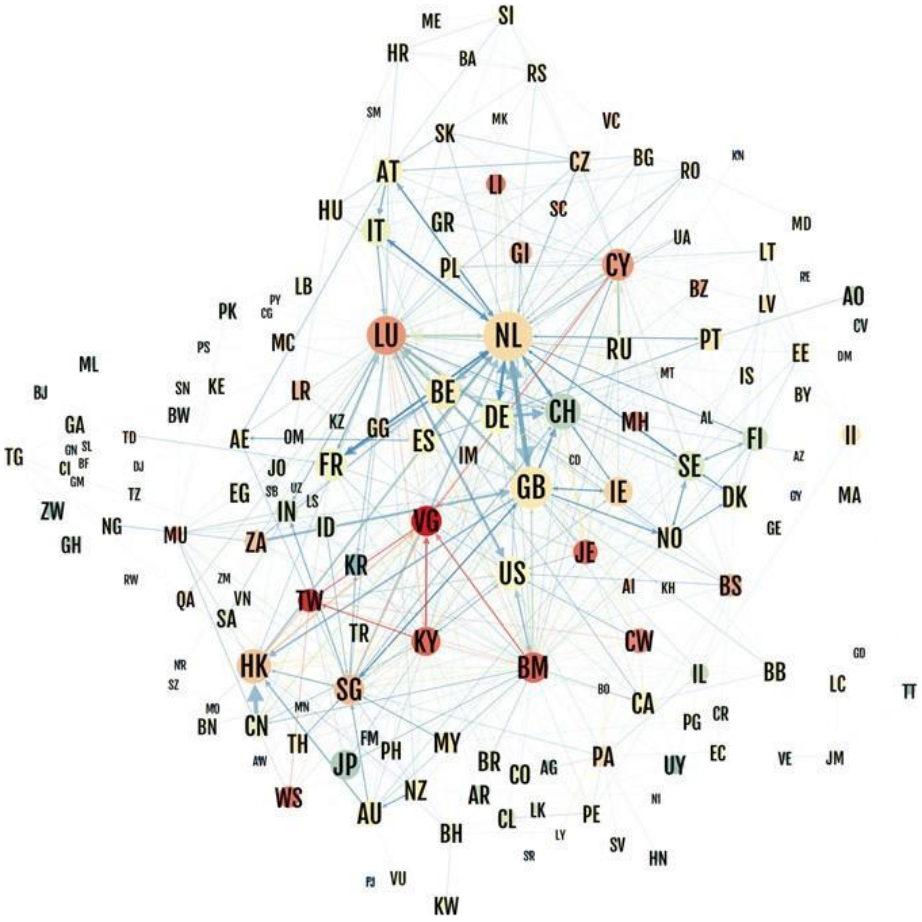
Financial data capture the annual financial information in financial reports published (not only for listed companies), such as turnover, profit rate or return on assets. The evidence is mixed on whether corrupt companies have good or bad financial performance compared to their clean peers. Posing a fundamental challenge for measurement, extraordinary financial performance such as high profit rates can also be driven by efficient companies and disruptive start-ups. Nevertheless, both quantitative and qualitative evidence shows that corruption tends to be related to odd financial performance of favoured companies, especially in the presence of other risk indicators such as public procurement tendering risks (Fazekas&Tóth, 2017).

Proxy indicators based on ownership and management data are possibly the most widely used for identifying company corruption risks. Companies used for corrupt ends typically hide final beneficial owners: either by registering in a tax haven (de Willebois et al., 2011) (Figure 3) or using complex ownership structures or both (Garcia-Bernardo et al., 2017). Alternatively, strawmen are used to avoid public exposure that often leads to odd management profiles, e.g. using one acting manager/director for multiple companies (Jancsics&Jávor, 2012). Opaque corporate structures often used to carry out the transaction, but they can also be used to hide and secure the proceeds from corruption (Jancsics, 2018). Due to a global push for publishing beneficial ownership information, especially in specific sectors such as extractives, some countries start to provide more transparent and accessible ownership information, albeit data

quality and accuracy remains a major issue.[7] While detailed management and ownership data is probably one of the most useful ways to track company risks, it is also the least widely available and hardest to work with (Heemskerk et al., 2018).

Overall, company risk indicators hold great promise, especially considering the increasing availability of company data. Similar to procurement indicators, they allow analysing corruption both at the micro-level (e.g. organisations) and across countries, as such data is often published in standardized ways. However, these indicators are seldom tested and differ widely in the kinds of risks they measure.

Figure 3. Network of ownership flows between countries, with high-risk jurisdictions highlighted in red



Source: (Garcia-Bernardo et al., 2017, fig. 3)

3.3.4 Personal connections

Personal connections between public officials and companies are one of the most widely used and oldest corruption risk indicators in government contracting as well as other government functions (Fisman, 2001). These connections are of diverse nature encompassing a range of direct and indirect connections traced:

- informal personal ties such as family or belonging to the same professional association or sports club (Grodeland, 2005; Mungiu-Pippidi, 2011),

[7] https://www.openownership.org/

- revolving door (Amore&Bennedsen, 2013; Bó, 2006; Dombrovsky, 2008),
- brokers and intermediary organisations establishing personal links (Rajwani&Liedong, 2015).
- simultaneously holding public and private offices (Etzioni, 2009),
- political party and campaign contributions (Boas et al., 2014; Bromberg, 2014; Fazekas et al., 2022; OECD, 2017; Witko, 2013), and
- lobbying (Dávid-Barrett, 2011).

The use of these different strategies of personal connections and the ways in which they are combined depend on the threat of exposing corruption and the legal framework (e.g. conflict of interest regulations) (Trapnell, 2011). Theoretically, these different forms of political connections, personal or impersonal, direct or indirect, are expected to work similarly: first, political ties represent a means of controlling and managing the transaction in an informal, corrupt contract which is typically non-enforceable by courts. Second, they also serve as a vehicle for rent extraction when the political officeholder earns income from a linked company. Third, political connections can also support broader trust building and facilitate information sharing, especially when the corrupt network is large and the benefits and costs of corruption are spread across the network.

However, private-public sector ties may not only signal higher corruption risks, but also a host of benefits such as improved information flows, the spread of entrepreneurial values, etc. Hence, it is only the misuse of connections which poses risks of corruption. As such high-quality indicators not only measure the existence of a particularistic link, but also measure its impact. For example, in public procurement, many studies look at the amount of contracts won due to connections while others link contracting processes and administrative structures to the presence of connections (Fazekas et al., 2022). For example, in Brazil, suppliers' political contributions result in additional contracts won worth 14 times more than the contributions (Boas et al., 2014), the same ratio in the US is only 2.5 times (Bromberg, 2014). Moreover, in the US, the strongest predictor of company contract volume from before to after the 1994 government change is the party to which the company was connected (Goldman et al., 2013). Surprisingly, even in Denmark -- one of the least corrupt countries globally -- direct family ties between companies and politicians increase company profitability, especially in sectors dependent on public procurement (Amore&Bennedsen, 2013).

In sum, personal connections indicators have high accuracy, but they are most useful when combined with other data, typically transactional. Indeed, the existence of a connection only shows heightened *opportunities* for abuse. These indicators are most useful when studying the social structure underpinning corrupt transactions and group behaviour.

4. DISCUSSION: HOW FAR DID WE GET? WHERE TO GO FROM HERE?

4.1 Navigating the measurement landscape

This review argues that measuring corruption is feasible if the right resources, analytical focus and sufficient perseverance is warranted. The dominance of expert, elite and population surveys in corruption measurement has gradually eroded in the last decade to give way to a more diverse landscape using a plethora of methods tailored to different types of corruption. Unsurprisingly, once one tries to collate corruption indicators relating to different levels of observation and corruption types, there is typically weak agreement among them. It is possible to have high corruption prevalence in one sector, say awarding government contracts, but having low prevalence in another, say obtaining individual healthcare services (Rose-Ackerman, 2015). Of course, disagreement between different indicators of corruption may indicate poor measurement quality, but only if the measured phenomena are closely aligned.

To help researchers and policy makers who need to find the right indicator for their specific goal, we offer an overview of indicator groups and their typical uses (Table 1). This table organises indicators along the lines of the above review looking at each main indicator group and offering a few widely used examples. We offer a few typical examples of indicator use and outline main pros and cons for using the indicators to the given aims. For example, a researcher or policy analyst who is interested in ranking countries in a region such as Eastern Europe would best look at expert ratings which are largely consistent across countries and easy to use and access. This would not inform changes in rankings over time, rather the overall position of countries relative to each other. Another, very different example is when an analyst is interested in the impact of specific policy changes such as tightening conflict of interest or transparency rules on corruption. In this case, the researcher or policy analyst should use one of the indicators which is sensitive to change and able to point at specific changes in impacted corrupt practices rather than high-level aggregate values. For example, public procurement data-based corruption risk indicators are often used in such impact evaluation studies.

Table 1. Overview table: diverse uses of different corruption indicators

Indicator group	Indicator Examples	Level of observation	Typical uses	Pros	Cons
Expert ratings	V-Dem CPI WGI	Country	-Cross-country comparisons -macro-determinants of corruption	-Wide coverage -Standardized data	-Noisy -Insensitive to change
Public opinion surveys	World Bank Enterprise Survey Barometers (e.g. Afro) Global Corruption Barometer	Individual or country	-Assessing perceptions and experiences of (petty) corruption -individual determinants of corruption (e.g. gender)	-Relatively wide coverage -Relatively standardized data	-They only measure specific experiences (e.g., bribery) -Perceptions of corruption are poor indicators of actual corruption
Proven cases	Transactional Records Access Clearinghouse (TRAC)	Country or region (state)	-tracking changes in corruption prevalence over time within a country or region	-Strong claim on validity -Relative consistency over time	-Only captures narrowly defined corrupt behaviours
Public employment	Meritocratic recruitment and promotion	Individual or public organisation	-tracing the degree of political influence on public bureaucracies -assessing the impact of meritocracy on economic outcomes (e.g. growth)	-Rich detail -Some country coverage of survey data	-Survey measures suffer from reporting biases (e.g. social desirability) -administrative data is rarely publicly available
Proxy indicators: public procurement	Corruption Risk Index (CRI)	Government contract/tender	-Evaluating impact of targeted interventions -comparing corruption risk prevalence within country (e.g. across public bodies)	-Rich detail (e.g. individual contracts or markets) -Wide country coverage -Consistent time series	-Risk indicators typically suffer from false positives -Exceptions from contract publication requirements
Proxy indicators: personal connections	Campaign contributions	Company	-Impact of major political changes (e.g. government change) -Investigating the corruption-growth relationship	-Rich detail (e.g. individual company performance) -Consistent time series	-Difficulty of tracing all relevant connections (false negatives) -Lack of comparability across countries

4.2 Agenda for future research

Our starting point for a future research agenda aiming to advance the corruption measurement landscape is that no single method is adequate on its own to measure all aspects and forms of corruption. Corruption is diverse, it manifests itself in many forms, only if we match measurement instruments to the type of corruption measured can we be successful. In particular, the time of surveys of corruption and detailed case studies has passed as the only solutions. They need to be complemented by more focused approaches typically drawing on Big ‘administrative’ Data which describes administrative behaviour at the micro, transactional level.

Understanding the key strengths of recent innovations, the proposed agenda shall further improve and refine corruption indicators so that they

- rest on a theoretically sound understanding of the specific corrupt transactions measured,
- derive from objective data describing actor behaviour,
- are defined on the transactional level (i.e. where corruption takes place),
- allow for consistent comparisons across countries, organisations, and time, and
- passed rigorous tests of measurement validity.

While corruption is secret, it usually leaves traces in official records such as public tenders, company ownership and financial information, or public sector hiring. As open access to, fair competition for, and transparency of public resources are prescribed by legal frameworks across every developed and in many developing countries, corruption, that is particularistic limitations of open access, must pretend that it is legal. This characteristic of corruption creates the opportunity for a range of indirect measurement approaches, following from anomalies of open market competition and fair distribution of public resources. In addition, the competition between corrupt groups and especially the change of power between them (e.g. which predatory elite group forms government) create a unique opportunity to identify what is impartial resource allocation and what is only a pretence of it (David-Barrett&Fazekas, 2019).

Some of the government functions where such an agenda is likely to bear fruits[8] include, but not limited to

- Making of laws and regulations,
- Public subsidies,
- Loans and guarantees for enterprises,
- Sale and rental of public assets,
- Award of public licenses,
- Public employment such as hiring and promotions, and
- Welfare payments.

[8] For novel indicators covering some of these fields such as integrity in law-making see the OECD's Public Integrity indicators: https://oecd-public-integrity-indicators.org/

BIBLIOGRAPHY

Amore, M. D., & Bennedsen, M. (2013). The value of local political connections in a low-corruption environment. *Journal of Financial Economics*, *110*(2), 387–402.

Auriol, E., Flochel, T., & Straub, S. (2011). *Public Procurement and Rent-Seeking: The Case of Paraguay* (No. 11–224; TSE Working Papers, Issues 11–224). Toulouse School of Economics (TSE).

Banerjee, A. V, & Pande, R. (2007). *Parochial Politics: Ethnic Preference and Politician Corruption* (2007.147; CID Working Paper Series).

Bó, E. D. (2006). Regulatory capture. A review. *Oxford Review of Economic Policy*, *22*(2), 203–225.

Boas, T. C., Hidalgo, F. D., & Richardson, N. P. (2014). The Spoils of Victory: Campaign Donations and Government Contracts in Brazil. *Journal of Politics*, *76*(2), 415–429.

Bromberg, D. (2014). Can Vendors Buy Influence? The Relationship Between Campaign Contributions and Government Contracts. *International Journal of Public Administration*, *37*(9), 556–567.

Campos, N., Engel, E., Fischer, R. D., & Galetovic, A. (2019). *Renegotiations and Corruption In Infrastructure: The Odebrecht Case* (No. 230; Marco Fanno Working Papers).

Caneppele, S., Calderoni, F., & Martocchia, S. (2009). Not only banks. Criminological Models on the Infiltration of Public Contracts by Italian Organized Crime. *Journal of Money Laundering Control*, *12*(2), 151–172.

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, Connections, and Corruption Risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. *Journal of Politics*, *79*(1), 89–103.

Charron, N., Dijkstra, L., & Lapuente, V. (2014). Regional Governance Matters: Quality of Government within European Union Member States. *Regional Studies*, *48*(1), 68–90.

Chong, E., Klien, M., & Saussier, S. (2015). *The Quality of Governance and the Use of Negotiated Procurement Procedures: Evidence from the European Union* (No. 2015–3; EPPP DP).

Christensen, J. (2011). The looting continues: tax havens and corruption. *Critical Perspectives on International Business*, *7*(2), 177–196.

Cordis, A. S., & Milyo, J. (2016). Measuring Public Corruption in the United States: Evidence From Administrative Records of Federal Prosecutions. *Public Integrity*, *18*, 127–148.

Dahlström, C., Fazekas, M. and Lewis, D.E. (2021), Partisan Procurement: Contracting with the United States Federal Government, 2003–2015. American Journal of Political Science, 65: 652-669.

Dahlström, C., Lapuente, V., & Teorell, J. (2012). The Merit of Meritocratization: Politics, Bureaucracy and the Institutional Deterrents of Corruption. *Political Research Quarterly*, *65*(3), 656–668.

Dávid-Barrett, E. (2011). *Cabs for Hire? Fixing the Revolving Door Between Government and Business*.

David-Barrett, E., & Fazekas, M. (2019). Grand corruption and government change: an analysis of partisan favoritism in public procurement. *European Journal on Criminal Policy and Research*.

De Vries, C. E., & Solaz, H. (2017). The Electoral Consequences of Corruption. *Annual Review of Political Science*, *20*(1), 391–408.

de Willebois, E. van der D., Halter, E. M., Harrison, R. A., Park, J. W., & Sharman, J. C. (2011). *The Puppet Masters: How the Corrupt Use Legal Structures to Hide Stolen Assets and What to Do About It*. World Bank Publications.

Decarolis, F., & Giorgiantonio, C. (2019). *Corruption Indicators in Public Procurement: Evidence from Italian Calls for Tenders*.

Dombrovsky, V. (2008). *Do Political Connections Matter? Firm-Level Evidence From Latvia* (No. 3; SSE Riga/BICEPS Research Papers).

Escresa, L., & Picci, L. (2015). A New Cross-National Measure of Corruption. *World Bank Economic Review*.

Etzioni, A. (2009). The Capture Theory of Regulations—Revisited. *Society*, *46*(4), 319–323.

Fazekas, M., Cingolani, L., & Tóth, B. (2018). Innovations in Objectively Measuring Corruption in Public Procurement. In H. K. Anheier, M. Haber, & M. A. Kayser (Eds.), *Governance Indicators. Approaches, Progress, Promise* (pp. 154–185). Oxford University Press.

Fazekas, M., Ferrali, R., & Wachs, J. (2022). Agency independence, campaign contributions, and favouritism in US federal government contracting. *Journal of Public Administration Research and Theory*, muac026.

Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data. *British Journal of Political Science*, *50*(1), 155–164.

Fazekas, M., & Tóth, B. (2017). *Proxy indicators for the corrupt misuse of corporations* (2017:6; U4 Brief).

Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal of Criminal Policy and Research*, *22*(3), 369–397.

Fazekas, M., & Wachs, J. (2020). Corruption and the Network Structure of Public Contracting Markets across Government Change. *Politics and Governance*, *8*(2), 153–166.

Ferraz, C., & Finan, F. (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics*, *123*(2), 703–745.

Ferwerda, J., Deleanu, I., & Unger, B. (2017). Corruption in Public Procurement : Finding the Right Indicators. *European Journal on Criminal Policy and Research*, *23*(2), 245–267.

Fisman, R. (2001). Estimating the value of political connections. *The American Economic Review*, *91*(4), 1095–1102.

Garcia-Bernardo, J., Fichtner, J., Takes, F. W., & Heemskerk, E. M. (2017). Uncovering Offshore Financial Centers: Conduits and Sinks in the Global Corporate Ownership Network. *Nature Scientific Reports*, *7*, 6246.

Gerardino, M. P., Litschig, S., & Pomeranz, D. (2017). *Can Audits Backfire? Evidence from Public Procurement in Chile* (No. 23978; NBER Working Paper).

Golden, M. A., & Picci, L. (2005). Proposal for a New Measure of Corruption, illustrated with Italian data. *Economics & Politics*, *17*(1), 37–75.

Goldman, E., Rocholl, J., & So, J. (2013). Politically Connected Boards of Directors and The Allocation of Procurement Contracts. *Review of Finance*, *17*(5), 1617–1648.

Gordon, S. C. (2011). Politicizing agency spending authority: Lessons from a Bush-era scandal. *American Political Science Review*, *105*(04), 717–734.

Grodeland, A. B. (2005). Bulgaria, Czech Republic, Romania and Slovenia: The Use Of Contacts and Informal Networks in Public Procurement. In OECD (Ed.), *Fighting Corruption and Promoting Integrity in Public Procurement* (pp. 59–76). OECD.

Heemskerk, E., Young, K., Takes, F. W., Cronin, B., Garcia-Bernardo, J., Henriksen, L. F., Winecoff, W. K., Popov, V., & Laurin-Lamothe, A. (2018). The promise and perils of using big data in the study of corporate networks: problems, diagnostics and fixes. *Global Networks*, *18*(1), 3–32.

Hollibaugh Jr., G. E., Horton, G., & Lewis, D. E. (2014). Presidents and Patronage. *American Journal of Political Science*, *58*(4), 1024–1042.

IMF. (2016). *Corruption : Costs and Mitigating Strategies* (No. 16/05; Staff Discussion Note). IMF, Washington

Jancsics, D. (2018). Shell Companies and Government Corruption. In *Global Encyclopedia of Public Administration, Public Policy, and Governance* (pp. 4–21). Springer, Cham.

Jancsics, D., & Jávor, I. (2012). Corrupt Governmental Networks. *International Public Management Journal*, *15*(1), 62–99.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2006). Measuring Corruption: Myths and realities.

Development Outreach.

Klasnja, M. (2016). Corruption and the Incumbency Disadvantage: Theory and Evidence. *Journal of Politics*, *77*(4), 928–942.

Lambsdorff, J. G. (2005). *The Methodology of the 2005 Corruption Perceptions Index*.

Lambsdorff, J. G. (2006). Measuring Corruption – The Validity and Precision of Subjective Indicators (CPI). In C. Sampford, A. Shacklock, C. Connors, & F. Galtun (Eds.), *Measuring Corruption* (pp. 81–100). Ashgate.

Lopez-Iturriaga, F. J., & Sanz, I. P. (2018). Predicting Public Corruption with Neural Networks: An Analysis of Spanish Provinces. *Social Indicators Research*, *140*(3), 975–998.

Maeda, K., & Ziegfeld, A. (2015). Socioeconomic status and corruption perceptions around the world. *Research & Politics*, *2*(2).

Mancini, P., Mazzoni, M., Cornia, A., & Marchetti, R. (2016). Representations of Corruption in the British, French, and Italian Press: Audience Segmentation and the Lack of Unanimously Shared Indignation. *The International Journal of Press/Politics*, *22*(1), 67–91.

Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics*, *110*(3), 681–712.

Meyer-Sahling, J.-H. (2009). *Sustainability of Civil Service Reforms in Central and Eastern Europe. Five Years After EU Accession* (Issue 44). OECD.

Mironov, M., & Zhuravskaya, E. V. (2016). Corruption in Procurement and the Political Cycle in Tunneling: Evidence from Financial Transactions Data. *American Economic Journal: Economic Policy*, *8*(2), 287–321.

Mungiu-Pippidi, A. (Ed.). (2011). *Contextual Choices in Fighting Corruption: Lessons Learned*. Norwegian Agency for Development Cooperation.

Mungiu-Pippidi, A., & Fazekas, M. (2020). How to define and measure corruption. In A. Mungiu-Pippidi & P. M. Heywood (Eds.), *A Research Agenda for Studies of Corruption* (pp. 7–26). Elgar.

North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and Social Orders. A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.

OECD. (2017). *Preventing Policy Capture. Integrity in Public Decision Making*. OECD.

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, *90*(4–5), 853–870.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public Economics*, *93*(7–8), 950–964.

Olken, B. A., & Pande, R. (2012). *Corruption in Developing Countries*.

Rajwani, T., & Liedong, T. A. (2015). Political activity and firm performance within nonmarket research: A review and international comparative assessment. *Journal of World Business*, *50*, 273–283.

Reinikka, R., & Svensson, J. (2004). Local Capture: Evidence From a Central Government Transfer Program in Uganda. *The Quarterly Journal of Economics*, *119*(2), 678–704.

Rose-Ackerman, S. (2008). Corruption. In C. K. Rowley & F. G. Schneider (Eds.), *Readings in Public Choice and Constitutional Political Economy* (pp. 551–566). Springer US.

Rose-Ackerman, S. (2015). Are Corrupt Elites Neccessary for Corrupt Countries? In C. Dahlström & L. Wangnerud (Eds.), *Elites, Institutions and the quality of government* (pp. 33–48). Palgrave Macmillan.

Rose-Ackerman, S., & Palifka, B. J. (2016). *Corruption and Government: Causes, Consequences, and Reform* (2nd ed.). Cambridge University Press.

Rose, R., & Peiffer, C. (2015). *Paying Bribes for Public Services. A Global Guide to Grass-Roots Corruption*. Palgrave Macmillan.

Rose, R., & Peiffer, C. (2019). *Bad Governance and Corruption*. Palgrave Macmillan.

Schuster, C., Meyer-Sahling, J.-H., & Mikkelsen, K. S. (2020). (Un)principled principals,

(un)principled agents: The differential effects of managerial civil service reforms on corruption in developing and OECD countries. *Governance*, *33*(4), 829–848.

Sequeira, S. (2012). Advances in Measuring Corruption in the Field. In D. Serra & L. Wantchekon (Eds.), *New Advances in Experimental Research on Corruption* (pp. 145–176). Emerald Publishing.

Sundell, A. (2014). *Nepotism and meritocracy* (2014:16; Working Paper Series).

Svensson, J. (2003). Who Must Pay Bribes and How Much? Evidence from a Cross Section of Firms. *The Quarterly Journal of Economics*, *118*(1), 207–230.

Svensson, J. (2005). Eight Questions about Corruption. *Journal of Economic Perspectives*, *19*(3), 19–42.

Transparency International. (2006). *Handbook for Curbing Corruption in Public Procurement*. Transparency International.

Transparency International. (2012). *Gateway: Mapping The Corruption Assessment Landscape*. Transparency International.

Trapnell, S. (2011). Actionable Governance Indicators: Turning Measurement into Reform. *Hague Journal on the Rule of Law*, *22*(3), 317–348.

Trapnell, S. (2015). *User's Guide to Measuring Corruption and Anti-Corruption*. United Nations Development Programme.

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*, 211–244.

Vaz Mondo, B. (2016). Measuring Political Corruption from Audit Results: A New Panel of Brazilian Municipalities. *European Journal of Criminal Policy and Research*, *22*(3), 477–498.

Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption risk in contracting markets: a network science perspective. *International Journal of Data Science and Analytics*, *12*(1), 45–60.

Witko, C. (2013). Campaign contributions, access, and government contracting. *Journal of Public Administration Research and Theory*, *23*(2), 457–494.

Appendix: Overview of Main corruption datasets

Table A1. Overview of selected corruption data sources

Name	Type	Content (coverage, period)	Link
Surveys	World Governance Indicators (WGI)	214 countries 1996-2020 (25 years)	http://info.worldbank.org/governance/wgi/
	Corruption Perception Index (CPI)	180 countries 1995-2021 (27 years)	https://www.transparency.org/en/cpi/2021
	Varieties of Democracy (VDEM)	202 countries 1789-2021 (203 years)	https://www.v-dem.net/vdemds.html
	European Quality of Government Index	30 countries (regional samples) 2010-2021 (12 years)	https://www.gu.se/en/quality-government/qog-data/data-downloads/european-quality-of-government-index
	EuroBarometer	134 countries 1999-2021 (22 years)	https://europa.eu/eurobarometer/surveys/browse/all
	AfroBarometer		http://www.afrobarometer.org
	AmericasBarometer		www.vanderbilt.edu/lapop
	ArabBarometer		https://www.arabbarometer.org/
	AsianBarometer		http://www.asianbarometer.org
	World Bank Enterprise surveys	153 countries 2002-2021 (20 years)	http://www.enterprisesurveys.org/
	Global Corruption Barometer	119 countries 2003-2017 (15 years)	https://www.transparency.org/en/gcb
US federal corruption convictions	Measuring proven cases	State-year-level aggregation US 1986-2014	(Cordis & Milyo, 2016) https://tracfed.syr.edu/index/index.php?layer=cri
Cross-border bribery dataset	Measuring proven cases	Country-level aggregation Global 2000-2014	https://sites.google.com/site/lucioxpicci/measure_corruption
Government Transparency Institute	Micro-level proxies: procurement	Global (mainly) 2008-2022	https://www.govtransparency.eu/gtis-global-government-contracts-database/
Civil Service Surveys	Micro-level perceptions and proxies	Global Varying, recent years	https://www.globalsurveyofpublicservants.org/