An Axiomatic Approach to the Measurement of Corruption: Theory and Applications

James E. Foster, Andrew W. Horowitz, and Fabio Méndez

No generally accepted framework exists for constructing and evaluating measures of corruption. This article shows how the axiomatic approach of the poverty and inequality literature can be applied to the measurement of corruption. A conceptual framework for organizing corruption data is developed, and three aggregate corruption measures consistent with axiomatic requirements are proposed. The article also provides guidelines for empirical applications of corruption measures and discusses data requirements. A brief empirical example illustrates how each of the measures captures a distinct view of corruption that yields a different ranking. To the authors' knowledge, this article provides the first analysis of corruption measurement using an axiomatic framework. JEL codes: K42, O17, P37

Several recent articles have identified new data sources that allow the measurement of corruption to be based on actual episodes rather than perceptions of corruption. Seligson (2006) uses victim surveys to obtain quantitative data on the prevalence of bribery. Reinikka and Svensson (2006) use public expenditure tracking surveys to quantify embezzlement of public funds and enterprise surveys to quantify bribery at the micro level. Olken (2009) and Ferraz and Finan (2008) rely on external audits to measure fraud in local governments. Gorodnichenko and Sabirianova-Peter (2007) use gaps between the incomes and consumption of public officials for similar purposes. These evolving data sources will allow researchers and policymakers to pose new questions about corruption and create targeted policies to address it.

James E. Foster (fosterje@gwu.edu) is a professor of economics and international affairs at the George Washington University and Research Associate at OPHI, Oxford University. Andrew W. Horowitz (horowitz@uark.edu) is a professor of economics in the Sam M. Walton College of Business at the University of Arkansas. Fabio Méndez (corresponding author; fmendez@uark.edu) is an associate professor of economics in the Sam M. Walton College of Business at the University of Arkansas. The authors are grateful to participants at the Latin American Econometric Society–Latin American and Caribbean Economic Association conference in 2008, the Brazilian Econometric Society meeting in 2008, the Pontifical Catholic University of Rio de Janeiro, Midwest Economic Association meetings, and the Walton College of Business. Foster gratefully acknowledges research support provided by the Institute for International Economic Policy (IIEP) of the George Washington University.

Advance Access Publication March 1, 2012
© The Author 2012. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
But improved data on corruption will not automatically resolve questions or
guide policy. There are many ways of translating data into corruption mea-
sures, and results from any study could depend on the measurement lens that is
being used. Moreover, corruption is defined differently in different contexts
(Bardhan 2006); this inconsistency of definition appears in both empirical and
theoretical applications. It would seem to be an appropriate moment for ex-
ploring how corruption might be measured.

This article presents a framework for assessing and comparing measures of
corruption, enabling more effective use of existing data and providing guidance
for data collection. It adapts the axiomatic structure of poverty and inequality
measurement to organize corruption data and generate aggregate corruption
measures. The axiomatic approach involves defining potentially important
properties of corruption measures—axioms—and classifying the measures
based on those properties. Specifying the axiomatic properties of corruption
measures provides criteria that researchers and policymakers can use to evalu-
ate and classify corruption measures and interpret empirical findings.

The article is organized as follows. Section I introduces the general concep-
tual framework and terminology. Section II discusses axioms that could plaus-
ibly form the foundations for measuring corruption and shows that some
common measures are incompatible with those axioms. Section III presents
corruption measures that are compatible with the basic axioms and develops
additional properties that classify them further. Section IV defines data require-
ments for the proposed corruption measures and provides an illustrative
example using data from the Business Environment and Enterprise
Performance Survey developed by the World Bank and European Bank for
Reconstruction and Development. Section V concludes and suggests directions
for future research.

I. Terminology and Conceptual Framework

There are two types of individuals in the model: officials and clients. Officials
are public servants who perform functions such as selling government goods or
services, allocating government transfers and funds, issuing permits or levying
penalties related to government regulations, and similar tasks. All such func-
tions are called services, and the set of public officials associated with a specific

1. See Méndez and Sepúlveda (2010) for a theoretical model that illustrates the difficulties of using
alternative corruption measures to resolve a question.
2. For additional examples, see Wolfers (2006), Svensson (2003), and Clarke and Xu (2004), who
all use as their measure the number of corrupt transactions. In contrast, Olken (2009), Gorodnichenko
and Sabirianova-Peter (2007), Shleifer and Vishny (1993), and Choi and Thum (2005) use the amount
of money involved in corrupt transactions. And Cadot (1987) and Çule and Fulton (2005) use the
percentage of government officials who are party to corrupt transactions.
3. Related research, including research that utilizes the survey, is available at www.worldbank.org/
wb/governance/pubs_statecapture.
service is called a department. Clients are members of the public who conduct business and use public services directly or indirectly. Services and departments are indexed by $s$ and clients by $i$. The total number of departments is defined as $S$ and the number of clients as $I$.

Transactions between departments and clients vary by purpose and size depending on the type of service provided. Examples include a legal payment for a passport application, a bribe paid for a driver license, and illegal appropriation of public funds allocated to a department. The analysis focuses on transactions between departments and clients because it is easier to obtain data on corrupt dealings by departments than by individual officials. Still, because the methodology treats officials and departments the same way, it could be applied to individual officials if the data were available.

Transactions are recorded in a $T \times I \times S$ data array $D$ containing $T$ many transaction reports, denoted by $d_t$. The data array has two types of entries: $d_{tis}$ can be the monetary value of a transaction observed between client $i$ and department $s$ in report $d_t$, or it can be an empty cell—indicating that no transaction was observed between $i$ and $s$ in report $d_t$. Every transaction report $d_t$ contains at least one transaction among its $S \times I$ many cells, but can also list several transactions between different clients and departments. Multiple transactions between a specific client and department are recorded in different reports. In general, $T$ and $t$ refer to numbers of transaction reports, not to specific periods of time.

As an example, consider the $2 \times 4 \times 4$ array $D$ in figure 1. This array consists of two reports, $d_1$ and $d_2$, each covering four clients and four departments. The entry $d_{113} = 7$ indicates that a transaction with a value of 7 is recorded in report 1 between client 1 and department 3. Other entries are missing—so, for example, $d_{111}$ indicates that report 1 records no transaction between department 1 and client 1. Also note that the transaction amount between a client and department might be zero. For example, a police officer might provide protection services to clients without receiving any direct payments from them.

Two additional aspects of the above framework are noted. First, the proposed framework does not track transactions between private parties. The

---

4. It is natural for each transaction report $d_t$ to have only one transaction. The examples use numerous transactions for each report for illustrative purposes.
focus is on corruption involving public servants. Second, the framework is
designed to record actual transactions, not implicit or counterfactual ones.
Thus it does not address passive corrupt acts such as bribe offers (or demands)
that are not accepted. Though such passive acts would ideally be captured by
an aggregate measure of corruption, data limitations may make it impractical
to assess.

II. An Axiomatic Approach to Measuring Corruption

The axiomatic approach to measuring poverty was pioneered by Sen (1976,
1983) with early application by Foster, Greer, and Thorbecke (1984). The
methodology involves two steps: identification and aggregation. When measur-
ing poverty, identification determines who is poor and aggregation maps the
data into an overall level of poverty. Similarly, measuring corruption requires
explicit identification criteria for determining when a particular transaction is a
corrupt transaction and a method of aggregating the resulting data into an
overall, scalar measure of corruption.

Identifying poverty typically involves specifying a poverty line based on
income or consumption, or on dimension-specific cutoffs in capability, where
persons having attainments below the poverty line are considered poor.
Analogously, a transaction can be identified as corrupt if the payment received
for the government service exceeds an allowable threshold. The question then
becomes: what is the appropriate threshold?

In some circumstances, it may be sensible to set the threshold at the legal
price of the service, so that a transaction is considered to be prima facie
corrupt if the payment to the official exceeds this price by any margin. In
general, though, identifying a corrupt transaction might not be so straight-
forward. For example, in countries it is acceptable for officials to accept a
gift as long as it can be consumed in one day. In others it is acceptable to
give cookies to government clerks on their birthdays—a practice that could
be indistinguishable from a small bribe. Whether the payment received or
the preferential treatment granted warrants defining a transaction as corrupt
depends on several factors including the legal price of the service, local
culture and habits, type of service provided, and institutional framework.\(^5\)

Each context could lead to a different threshold representing the associated
tolerance for excess payments.

This paper assumes that thresholds vary across services or departments,
but are otherwise fixed, and hence can be represented by a vector \(Z\) of
tolerance thresholds, one for each department. Threshold value \(z_s\) is the

\(^5\) Such examples bear some analogy to Sen’s (1976, 1983) discussion of the absolute and relative
 nature of poverty.
corrupt. Thus, transaction $d_{tis}$ from $D$ is considered corrupt if $d_{tis} > z_s$ but not if $d_{tis} \leq z_s$.\(^6\)

Using thresholds to identify a corrupt transaction means that only the total amount paid is needed to construct corruption measures. So, when gathering information from clients, investigators do not have to ask how much clients spent on bribes or illegal payments—just how much they spent on specific services (a question that the clients might be more willing to answer truthfully). This approach also allows researchers to use tolerance thresholds of zero ($z_s = 0$) to cover cases where clients had to pay for services that they should have received for free.

Aggregation is the next step in constructing a corruption measure. This step maps transactions in $D$ and thresholds in $Z$ into an aggregate level of corruption. Additional information on client income levels, as contained in a client resource vector $Y$ with individual levels $y_i$, might also be used in the aggregation process. Consequently, in this paper, a corruption measure is represented as a scalar valued mapping $C(D; Z, Y)$. The next question considered is: what are the basic properties that a corruption measure should exhibit?

**Axioms for Measuring Corruption**

The first task in applying the axiomatic approach is to specify a set of properties for corruption measures. Existing and proposed measures can then be evaluated and classified based on the properties they satisfy. This section begins with a set of basic axioms that all corruption measures might be expected to satisfy. Admittedly, any set of basic axioms could be challenged as being too exacting or too lenient. A primary goal of this article is to foster debate that might lead to consensus on a set of basic axioms for measuring corruption.

Consider a generic corruption measure $C(D; Z, Y)$ that uses tolerance vector $Z$ to convert the data in $D$ and client resources vector $Y$ into a corruption metric. In addition, assume that all transactions are expressed in the same real monetary units, eliminating concerns about inflation or units of measure. Four definitions lay the groundwork for stating the basic axioms:

- $D'$ is obtained from $D$ by a *reordering of client observations* if, for some pair $j$ and $k$ of distinct clients, the following holds for all $t$ and $s$: $d'_{tis} = d_{tks}$ and $d'_{iks} = d_{tjs}$, while $d'_{tis} = d_{tis}$ for all $i \neq j, k$.

In other words, all observations are the same except for two rows ($j, k$) whose elements have been switched on all $T$ records.

- $D'$ is obtained from $D$ by a *replication of observations* if there is an integer $m \geq 2$ such that $T' = mT$ and $D' = (D, \ldots, D)$ where $D'$ is an $mT \times I \times S$ array.

6. The identification threshold can be defined as the difference between a payment and $z_s$ to accommodate cases where a client (such as a relative) pays less than the standard price for a service.
In other words, $D'$ has $m$ copies of the records in $D$.

- $D'$ is obtained from $D$ by an increment if $d'_{tis} > d_{tis}$ for a given index combination $(t, i, s)$ and $d'_{ujr} = d_{ujr}$ for all $(u, j, r) \neq (t, i, s)$. The increment is within tolerance if $z_s \geq d'_{tis}$. It is frequency increasing if $d'_{tis} > z_s$. It is excess payment increasing if $d_{tis} > z_s$.

An increment occurs when a single payment is increased and all other entries are unchanged. If the payment begins and ends below the tolerance threshold, it is a within tolerance increment. If the payment begins below the threshold and ends above (and hence is now considered a bribe), the increment is frequency increasing. If the payment begins and ends above the threshold, so the size of the bribe is increased, then it is an excess payment increasing increment.

- $(D'; Z', Y')$ is obtained from $(D; Z, Y)$ by a proportionate change if $(D'; Z', Y') = \alpha(D; Z, Y)$ for $\alpha > 0$.

A proportionate change scales up or down all observations, incomes, and thresholds by the same factor.

**Basic Axioms**

With those definitions in mind, a set of basic axioms can be defined that all corruption measures would be expected to satisfy:

- **Client anonymity.** If $D'$ is obtained from $D$ by a reordering of client observations, then $C(D'; Z, Y) = C(D; Z, Y)$.

Client anonymity assures that the client index number does not impact the corruption measure. That is, it ensures that the identity of the private agents involved in corrupt transactions does not affect the resulting corruption measure; in line with the stated goal of capturing official corruption only. Notice, however, that Client Anonymity does not preclude the use of departmental weights among different government services.

- **Replication invariance.** If $D'$ is obtained from $D$ by a replication of observations, then $C(D', Z, Y) = C(D, Z, Y)$.

Replication invariance ensures that the measure does not depend on the absolute number of corrupt transactions, but rather on the number of corrupt transaction relative to total transactions. That way, the measure does not treat environments with fewer transactions more favorably, but instead measures corrupt transactions relative to the total number of government services provided.

- **Focus.** If $D'$ is obtained from $D$ by a within-tolerance increment, then $C(D'; Z, Y) = C(D; Z, Y)$.

The focus axiom ensures that the measure is unresponsive to payment amounts for transactions not involving corruption.
• **Frequency monotonicity.** If $D'$ is obtained from $D$ by using a frequency-increasing increment, then $C(D'; Z, Y) > C(D; Z, Y)$.

The frequency monotonicity axiom requires a corruption measure to increase when the value of a transaction crosses the tolerance threshold.

This simple set of axioms disqualifies many potential measures of corruption. For example, a department headcount ratio that measures corruption as the percentage of departments (or officials) that had accepted at least one bribe would violate the frequency monotonicity axiom. Similarly, measures that merely sum all identifiable corruption—such as the amount of money paid for bribes or the number of transactions identified as corrupt—would violate the replication invariance axiom. If corruption were measured as all payments above the legal price, the focus axiom could be violated, while measures that weight the transactions of different clients differently would violate the client anonymity axiom.

**Supplementary Axioms**

This section introduces a set of supplementary axioms that may be desirable in certain contexts. The axioms can help distinguish among different corruption measures and define more clearly what each measure is capturing.

• **Bribery monotonicity.** If $D'$ is obtained from $D$ by an increasing excess payment increment, then $C(D'; Z, Y) > C(D; Z, Y)$.

The bribery monotonicity axiom ensures that the corruption measure increases when the value of a bribe increases.

• **Client enrichment.** If $D$ contains at least one transaction with positive excess value and $Y' >> Y$, then $C(D; Z, Y') < C(D; Z, Y)$.

The client enrichment axiom implies that measured corruption should fall if clients become unambiguously richer and the values of bribes remain the same.

• **Decomposability.** Let $D'$ be a data array with the same number of clients and departments as array $D$ and let $E = (D, D')$ be the array obtained by combining them. If $n(D'), n(D)$, and $n(E)$ are the respective numbers of (non-missing) transactions they contain, then $C(E; Z, Y) = [n(D')/n(E)] C(D'; Z, Y) + [n(D)/n(E)] C(D; Z, Y)$.

In some cases it is useful to decompose corruption measures by subsets of transactions. For example, policymakers might want to analyze corruption by regions or departments or based on client characteristics. The decomposability axiom requires the overall corruption level to be the weighted sum of subset corruption levels, where the weights are the shares of transactions in the subsets.

7. Foster (2009) and Alkire and Foster (2011) describe a similar issue in the measurement of chronic and multidimensional poverty.
• Scale invariance. If \((D'; Z', Y')\) is obtained from \((D; Z, Y)\) by a proportionate change, then \(C(D'; Z', Y') = C(D; Z, Y)\).

This property requires the measure to view corruption in relative terms, so that if all thresholds, transactions sizes and resource levels were doubled, the measured level of corruption would be unchanged.

III. PROPOSED CORRUPTION MEASURES

The previous section demonstrated that the application of even the most basic axioms disqualifies a number of corruption measures. This section introduces a family of corruption measures consistent with all the basic axioms and reasonably compatible with available cross-sectional data. The measures are defined by fixing a corruption function \(f(d_tis; z_s, y_i)\), which indicates the corruption level of a single transaction \(d_tis\) given the departmental tolerance threshold \(z_s\) and client resource level \(y_i\). This mapping identifies and evaluates corruption at the transaction level.

An associated corruption array \(D_f\) replaces each transaction \(d_tis\) with its associated corruption level \(f(d_tis; z_s, y_i)\) while leaving untouched the empty cells in array \(D\). A corruption measure \(C_f\) can then be constructed by taking the simple unweighted mean value of all the nonempty entries in the corruption array, or \(C_f(D; Z, Y) = \mu(D_f)\). In other words, \(C_f\) is the sum of corruption levels \(f(d_tis; z_s, y_i)\) for all transactions divided by the number of transactions. Three possible forms of the corruption function and their associated corruption measures are analyzed below.

Frequency Measure of Corruption

The first measure is based on the simplest corruption function, which takes a value of 1 when a transaction is corrupt and 0 when it is not. Formally, define \(f_1\) by

\[
f_1(d_tis; z_s, y_i) = \begin{cases} 1 & \text{if } d_tis > z_s \\ 0 & \text{if } d_tis \leq z_s \end{cases}
\]

and denote the associated corruption array by \(D_1\). The frequency measure of corruption \(C_1(D; Z, Y) = \mu(D_1)\) measures corruption as the fraction of transactions that are corrupt. \(C_1\) is clearly bounded between 0 and 1, with higher numbers indicative of greater prevalence; it is analogous to the simple head-count ratio from the poverty literature. As an example, consider the 2x4x4 array \(D\) in figure 1 and suppose that the threshold vector is \(Z = [\$0, \$0, \$2, \$0]\). The corruption

\[\begin{array}{cccc}
0 & 1 & 0 \\
1 & 0 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{array}\]

Source: Authors’ construction.
function \( f_1 \) is applied to the transactions in \( D \) to obtain the corruption array \( D_1 \) in figure 2, which contains 1 for every transaction higher than its respective cutoff and 0 for every transaction that is not. The number of corrupt transactions (8) is divided by the overall number of transactions (30), to obtain the frequency of corruption \( C_1 = 0.27 \).

**Excess Value Measure of Corruption**

One disadvantage of using frequency measures of corruption is that they do not account for the amount of resources captured by corruption, or the depth of corruption. For example, an economy where 1 in 10 transactions is corrupt has a frequency measure of \( C_1 = 0.1 \) regardless of whether the typical corrupt transaction involves a bribe of $10 or $1 million.

The depth of corruption can be incorporated into a measure by letting the corruption function be the excess value of a transaction. Define \( f_2(d_{tis}; z_s, y_t) = (d_{tis} - z_s) \) if \( d_{tis} > z_s \), and \( f_2(d_{tis}; z_s, y_t) = 0 \) if \( d_{tis} \leq z_s \), and let \( D_2 \) be the associated corruption array. Then the excess value measure of corruption \( C_2(D; Z, Y) = \mu(D_2) \) evaluates corruption as the extent to which the average transaction exceeds its threshold level. In other words, \( C_2 \) is the aggregate amount of money paid in bribes divided by the total number of transactions. This measure takes on nonnegative values and is analogous to a poverty gap measure.

In the example, the corruption function \( f_2 \) and the threshold vector \( Z = \{0, 0, 2, 0\} \) yield the corruption matrix \( D_2 \) given in figure 3, which contains the excess payments to government officials beyond the tolerance thresholds. \( C_2 \) is then the sum of the entries (20) divided by the number of entries (30) or \( C_2 = 0.66 \).

**Relative Burden Measure of Corruption**

\( C_2 \) measures the average depth of corruption, but does not take into account the varying resources of clients or the size of the economy. Arguably, a given-sized bribe imposes a larger burden on a client with fewer resources. Similarly, a country in which 10 percent of GDP is spent on corruption may be viewed as more corrupt than one that spends 2 percent, even if the total amount of bribes is the same. \( C_2 \) and \( C_1 \) cannot make this distinction.

**Figure 3.** Corruption Array \( D_2 \)

Source: Authors’ construction.
To account for the relative burden of corrupt transactions, corruption function $f_3$ measures excess payments relative to client resources:

$$f_3(dt_{is}; z_s, y_i) = \begin{cases} 
\frac{(dt_{is} - z_s)}{y_i} & \text{if } dt_{is} > z_s, \\
0 & \text{if } dt_{is} \leq z_s,
\end{cases}$$

$D_3$ is the associated corruption array. The relative burden measure of corruption $C_3$ is defined as $C_3(D; Z, Y) = \mu(D_3)$, or the sum of the relative burdens divided by the total number of transactions. The calculation of this measure is similar to the previous examples, and so a separate illustration is omitted. Note that under reasonable assumptions $C_3$ is bounded between 0 and 1.

**Weighted Measures of Corruption**

The measures presented so far implicitly consider each department to be equally important for measuring corruption. Thus it does not matter for $C_1$, $C_2$, or $C_3$ whether corruption occurs in the department guarding nuclear materials or in a public library. Arguably, corruption measures should have the potential of weighting some departments more heavily than others based on their relative importance in a country’s institutional hierarchy.\(^{8}\)

Departments can be differentially weighted using a vector $w$ that assigns weights $w_s$ based on criteria developed by researchers or policymakers. Weights might be determined by subjective evaluations; or by objective indicators, such as the percentage of government workers in a department, the percentage of the government budget allocated to a department, or the percentage of government transactions processed by a department. Each measure developed in this section can be modified to incorporate weights during aggregation as follows: let $C_{kw}(D; Z, Y) = \mu_w(D_k)$, where $\mu_w$ is the weighted mean associated with vector $w$.

Table 1 indicates the basic and supplementary axioms satisfied by the proposed measures $C_1$, $C_2$, $C_3$, and their weighted counterparts. Each of the measures satisfies the basic axioms. $C_2$ and $C_3$ take the size of bribes into account and thus satisfy the bribery monotonicity axiom, while $C_1$ ignores the extent of excess payments and violates it. Neither $C_1$ nor $C_2$ takes into account client incomes, so both violate the client enrichment axiom, while $C_3$ expresses excess payments as a percentage of client income and so satisfies that axiom. Because $C_1$, $C_2$, and $C_3$ are aggregated as means, they all satisfy the decomposability axiom, and can all be used for analysis.\(^{9}\) Finally, $C_1$ and $C_2$ satisfy the scale invariance axiom while $C_2$ does not. The weighted versions $C_{1w}$, $C_{2w}$, and $C_{3w}$ satisfy the same properties as their unweighted counterparts.

---

8. See the analogous discussion in Alkire and Foster (2011). It may also make sense to use different weights for different classes of clients, though that is not done here.

9. This feature suggests that each measure is subgroup consistent. For example, if corruption in each region of a country falls, then this must be reflected in a lower national corruption level. See Foster and Sen (1997) for a related discussion in the context of measuring poverty and inequality.
This section applies data (described below) to estimate the measures of corruption—$C_1$, $C_2$, and $C_3$—developed in the previous section and to show how divergent results from specific measures can be linked to their underlying axiomatic properties. The goal is to show how a measure’s underlying properties determine the results that it yields, not to defend the country corruption rankings obtained.

When considering these three measures, two questions must be addressed. First, what data are required to construct these measures? Second, does each measure have independent value, or do they all correlate regardless of which axioms they satisfy?

### Data Requirements

The framework proposed in this article and the measures defined in the previous section require basic data on individual interactions between clients (the public, including representatives of firms) seeking services from government officials or departments (again, used interchangeably in this analysis), and the government officials associated with those services. These data include the number of interactions that occurred in a specific period, the amounts paid by clients (if any), and the services involved.

This type of information can be obtained through surveys of the public and of company managers, such as those currently conducted by the World Bank’s Enterprise Surveys and the Latin American Public Opinion Project (Seligson 2006; Reinikka and Svensson 2006). These surveys measure corruption based on personal experience, but the questions are phrased so that respondents can avoid potential self-incrimination. For example, a survey question might ask: “How often do firms like yours have to make unofficial payments to public officials to obtain permits?” This type of question makes it possible to obtain a reasonable proxy for the share of transactions with excess payments. The
framework proposed in this article, however, makes clear that it is also desirable for these surveys to collect information on the total number of transactions and interactions with the public service provider.

Beyond that basic level, one could also incorporate data on graft and embezzlement of public funds. While such data are difficult to obtain, public expenditure tracking surveys are a good source of information. These surveys are designed to track flows of resources in bureaucracies and thus are ideal for identifying and quantifying political and bureaucratic capture and leakages of funds (Reinikka and Svensson 2006). In addition, external audits can be used to measure the extent of fraud in local governments (Olken 2009; Ferraz and Finan 2008), as can gaps between the incomes and consumption of public officials (Gorodnichenko and Sabiriano-Peter 2007).

Information on graft, embezzlement, and fraud could easily be incorporated into a data array like $D$ (see figure 1). These types of corruption do not directly involve specific clients but can still be accounted for in the proposed framework by adding a state auditor to the client vector. A similar approach could be used to incorporate information about corruption uncovered by criminal investigations. As data collection on corruption improves, the quality and variety of the data that can be used to create arrays like figure 1 should expand.

Are Measures of Corruption Correlated?

Data from the Business Environment and Enterprise Performance Survey (World Bank and European Bank for Reconstruction and Development 2000) were used to determine whether measures of corruption correlate with one another. The survey covered 4,000 firms in 26 Central and Eastern European countries in 1999–2000, examining a wide range of interactions between firms and government. Based on interviews with firm managers and owners, the survey is designed to generate comparative measurements on topics such as corruption, state capture, lobbying, and the quality of the business environment. The data can then be linked to specific firm characteristics and performance.

Information from this survey was used to calculate values for $C_1$, $C_2$, and $C_3$. Data from the survey are not ideal for constructing these measures, but the survey did make it possible to reconstruct important elements of the $D$ data array—including information on the frequency and monetary value of corrupt acts. For example, question 28 of the survey asks how often firms like the respondents’ need to make unofficial payments to public officials in relation to seven government functions. These functions are: to get connected to utilities like electricity and telephony, obtain licenses and permits, deal with taxes, be awarded government contracts, manage customs and importing requirements, deal with courts, and influence laws and regulations. These functions are treated as the seven departments of the complete $D$ arrays.

In question 28, respondents were asked to estimate how often they had to make unofficial payments (always, mostly, frequently, sometimes, seldom, and never). For the purposes of this example, numerical values were assigned to
these answers: 100 percent to always, 80 percent to mostly, 60 percent to frequently, 40 percent to sometimes, 20 percent to seldom, and 0 for never.

Similarly, question 27 asked businesses what percentage of revenues (on average) firms like theirs typically pay annually in unofficial payments to public officials. Possible answers were 0, less than 1 percent, 1–1.99 percent, 2–9.99 percent, 10–12 percent, 13–25 percent, and more than 25 percent. The answers were given numerical values equal to the medians of each group of ranges (except for more than 25 percent, which was capped at 26 percent).

Together with question 27, questions 29 and 51 provide greater details about the magnitude of unofficial payments. Question 51 asked respondents to estimate their firm’s annual sales, assets, and debt to the nearest range, with possible answers ranging from less than $250,000 to $500 million or more. Again, the medians of these ranges were used. In turn, question 29 asked respondents to estimate the share of unofficial payments made at the different government departments. Combining the information from these questions made it possible to estimate the total amount spent on bribes at each department.

Finally, additional information regarding firms’ behaviors and perceptions was incorporated. In particular, question 24 was used, which asked firms what percentage of senior managers’ time was spent dealing with government officials about the application and interpretation of laws and regulations. Other questions asked respondents to report on the likelihood of finding an honest official, the predictability of public policies, and how much of an obstacle corruption was to doing business.

To derive the proposed corruption measures, information on the number of all transactions for specific pairs of clients and departments is needed. This information is not available in the dataset, so it was assumed to be the same for all pairs of clients and departments. The dataset reports how often corrupt payments were made on average, so by averaging these values for all respondents and departments, it was possible to obtain the measure $C_1$. A similar process was used to construct $C_2$ and $C_3$. In the case of $C_3$, each surveyed firm has reported the total excess payments as a percentage of total revenues and the percent of total excess payments going to each department. From this, the excess payment to department $s$ as a share of firm $i$’s revenue can be obtained, which is interpreted as $\sum_t (d_{itis} - z_s)/y_i$. The mean value of these aggregate relative burdens is used as the final estimate of $C_3$. The process for computing $C_2$ is identical to that for $C_3$, except that total payments $\sum_t (d_{itis} - z_s)$ are used.

10. The $C_2$ and $C_3$ measures described in the previous section use the mean over all transactions. The empirically constructed values in this section use the mean over $I \times S$ aggregates. This simplification does not cause any loss of generality because it was assumed that each pair of clients and departments had the same number of transactions. The calculated values for $C_2$ and $C_3$ use a constant (the number of transactions per pair) multiplied by the original values and so preserve the rankings.
The orderings that result from the calculations of the three corruption measures were then compared. As a benchmark, the comparison used Transparency International’s Corruption Perceptions Index for 1999 and 2000. (For two countries the index was not available for 2000.) Perception indexes are often used to assess aggregate bureaucratic corruption and so offer a natural comparison to the three proposed measures. The index ranges from 0 to 10, with a higher number indicating less corruption. The analysis here inverted the index, so a higher number indicates more corruption.

A Spearman rank correlation matrix of the resulting country rankings is shown in table 2, with countries ranked from the most to least corrupt. Country rankings were used because the scales of the three proposed measures are not directly comparable, though similar conclusions are obtained if their levels are used instead. The table shows a positive and significant rank correlation between $C_1$ (corruption frequency measure) and $C_3$ (relative burden of corruption), and a negative but insignificant rank correlation between $C_1$ and $C_2$ (absolute costs of corruption measure).

Rankings from the Corruption Perceptions Index are positively and significantly correlated with those of $C_1$ and $C_3$ but negatively and significantly correlated with those of $C_2$. Given that $C_2$ is the only proposed measure that violates the scale invariance axiom, the negative correlations between $C_2$ and all the other measures might be associated with its scale properties. In addition, the positive correlations between $C_1$, $C_3$, and the Corruption Perceptions Index suggest that corruption perceptions may be processed independently of scale. More generally, table 2 suggests that the three measures provide different perspectives on corruption despite having only minor differences in axiomatic properties.

Complete rankings for the entire country sample are shown in table 3, which splits the sample into four groups based on their rankings on the Corruption Perceptions Index: low (0–5.9), lower middle (5.9–6.7), upper middle (6.7–7.6), and high (above 7.6). One aspect of the table is striking: The three measures reveal different corruption patterns for countries with similar rankings on the perceptions index. Consider Armenia and Romania. Both have upper-middle rankings on the Corruption Perceptions Index, but while Armenia is ranked low by $C_1$ and high by $C_3$, Romania is ranked high by $C_1$ and low by $C_3$. These findings suggest that although the Index’s measure of perceived corruption would rank both countries similarly, the types of corruption affecting these countries may differ. A detailed analysis of why corruption perceptions deviate from the findings of the axiom-based measures is beyond the scope of this article. The objective here is simply to note that these measures deviate significantly from each other and from perception-based measures.

11. Other frequently cited perception indexes include the International Country Risk Guide from Political Risk Services and the Institute for Management Development index of corruption. Both are closely correlated with the Corruption Perceptions Index.
and that axiomatic criteria can illuminate the underlying sources of these discrepancies.

But can individual measures provide specific insights on corruption? The findings here suggest that axiom-based measures might contribute to key issues in the literature. Consider the debate on whether corruption facilitates commerce by enabling businesses to circumvent bureaucratic delays or undermines it by weakening public institutions and worsening delays. Some authors have found support for the second hypothesis. Kauffman and Wei (1999) find a positive and significant relationship between firms’ perceptions of corruption and the amount of time they report wasting with bureaucracy (see Meon and Sekkat 2005 for further support of the second hypothesis). But because the authors used measures of perceived corruption, they cannot identify specific factors that shape managerial decisions about time allocation.

The analysis also finds a significant positive correlation between levels of perceived corruption and the time wasted by firms’ managers dealing with government officials, as shown in table 4. But a more detailed picture emerges when reviewing the correlation coefficients for the three measures. Time wasted has a positive and significant correlation with $C_2$ and $C_3$, and their coefficients are similar to that for the Corruption Perceptions Index. However, the correlation coefficient between time wasted and $C_1$ is only 0.07 and is not significantly different from zero.

Thus it appears that the prevalence or frequency of corruption has less influence on managerial time allocation decisions than do measures of the depth of corruption. In other words, that corruption related to frequent but petty processes—such as paying utility bills or complying with traffic regulations—may be less harmful than, say, corruption related to rigging contracts. And recall that $C_2$ and $C_3$ satisfy the bribery monotonicity axiom and $C_1$ does not.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Corruption frequency, $C_1$</th>
<th>Absolute costs of corruption, $C_2$</th>
<th>Relative costs of corruption, $C_3$</th>
<th>Corruption Perceptions Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>1</td>
<td>−0.27</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Absolute costs of corruption, $C_2$</td>
<td>0.52***</td>
<td>−0.26</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Relative costs of corruption, $C_3$</td>
<td>0.63***</td>
<td>−0.37*</td>
<td>0.67**</td>
<td></td>
</tr>
<tr>
<td>Corruption perceptions Index</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

** Significant at the 5 percent level. * Significant at the 10 percent level.

Source: Authors’ analysis is based on data from World Bank and EBRD (2000) and Transparency International using data from 1999 and 2000.
### Table 3. Rankings for Eastern and Central European Countries Based on Corruption Measures, 1999–2000

<table>
<thead>
<tr>
<th>Groupa</th>
<th>Country</th>
<th>Corruption frequency, C1</th>
<th>Absolute costs of corruption, C2</th>
<th>Relative costs of corruption, C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (0–5.9)</td>
<td>Slovenia</td>
<td>3</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Estonia</td>
<td>4</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Hungary</td>
<td>2</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Belarus</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Poland</td>
<td>13</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Lithuania</td>
<td>17</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Lower middle (6.0–6.7)</td>
<td>Latvia</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Croatia</td>
<td>6</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Bosnia and Herzegovina</td>
<td>14</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Slovakia</td>
<td>15</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Czech Republic</td>
<td>7</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Turkey</td>
<td>21</td>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Macedonia, FYR</td>
<td>23</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Bulgaria</td>
<td>16</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Upper middle (6.8–7.6)</td>
<td>Kazakhstan</td>
<td>10</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Uzbekistan</td>
<td>20</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Romania</td>
<td>24</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Moldova</td>
<td>19</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Armenia</td>
<td>9</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>High (7.7–10.0)</td>
<td>Russian Federation</td>
<td>11</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Albania</td>
<td>25</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Ukraine</td>
<td>22</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Georgia</td>
<td>18</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Azerbaijan</td>
<td>26</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Kyrgyz Republic</td>
<td>12</td>
<td>3</td>
<td>22</td>
</tr>
</tbody>
</table>

*Source:* Authors’ analysis based on data from Transparency International using data from 1999 and 2000.

*Note:* Groups are defined based on their rankings in Transparency International’s Corruption Perceptions Index.

* The Corruption Perceptions Index, from 0 to 10, is inverted in this table, so a higher number indicates a greater perception of corruption.

### Table 4. Correlations between Business Variables and Corruption Measures and Perceptions in Eastern and Central Europe

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Corruption frequency, C1</th>
<th>Relative burden of corruption, C2</th>
<th>Absolute costs of corruption, C3</th>
<th>Corruption Perceptions Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time wasted on bureaucracy</td>
<td>0.07</td>
<td>0.39*</td>
<td>0.34*</td>
<td>0.36*</td>
</tr>
<tr>
<td>Investment</td>
<td>−0.35*</td>
<td>0.02</td>
<td>−0.52*</td>
<td>−0.55*</td>
</tr>
</tbody>
</table>

* Significant at the 10 percent level.

Another issue often addressed in the literature is whether corruption hinders investment and therefore growth. Mauro (1995), for example, reports a negative relationship between aggregate investment levels and aggregated indexes of corruption perceptions. In the sample, firm managers were asked to estimate how much their investments had increased over the previous three years. The correlations between their answers and the three corruption measures, as well as the Corruption Perceptions Index, are shown in table 4.

The results confirm a negative correlation between investment and corruption perceptions but provide a nuanced assessment of that relationship. The relative burden of corruption, $C_3$, has a negative and significant correlation with investment, and the magnitude is about the same as that for the Corruption Perception Index. The corruption frequency measure, $C_1$, has a smaller (though still significant) correlation with investment decisions than does the index. The absolute costs of corruption measure, $C_2$, has essentially no correlation with investment. The implication is that given two otherwise identical countries with the same Corruption Perceptions Index rating, the country with a higher $C_3$ will experience less investment. Such analyses hold the promise of improving our understanding of the corruption-growth relation, by identifying the specific aspects of corruption that hinder investment, and our understanding of properties of the measures that yield these results.

V. Conclusion

To the authors’ knowledge, this article is the first application of an axiomatic framework to corruption measurement. The main goal is to initiate debate on the explicit properties that are desirable when measuring corruption. To this end, four such properties—the basic axioms—were proposed: client anonymity, replication invariance, focus, and frequency monotonicity. In addition, four supplementary axioms specify properties that may be desirable in specific contexts: bribery monotonicity, client enrichment, decomposability, and scale invariance. The article proposed three measures of corruption and classified them based on their axiomatic properties.

Available data do not permit exact calculations of these measures, but they do allow approximations. These approximations revealed significant discrepancies between measures with distinct axiomatic properties. Indeed, the empirical exercise in the previous section showed that reasonable corruption measures with distinct axiomatic properties may exhibit negative correlations. Such discrepancies highlight the fact that corruption is a multidimensional phenomenon that may be plausibly measured in several distinct ways.

These results call for theoretical and empirical researchers alike to clearly specify their definitions of corruption, the measurement criteria associated with those definitions, and the robustness of their results using alternative measures—because their findings may depend on the definition of corruption used. As the inventory of corruption data expands, our framework for
organizing the data, constructing corruption measures and assessing their axiomatic properties, provides criteria for evaluating alternative measures that do not yet exist. In addition, the framework in this article suggests additional survey questions that can make corruption measures more useful and comparable.

The corruption measures in this article are defined for a given period. They do not focus on trends in corruption for specific clients or departments over time. For example, the data from the Business Environment and Enterprise Performance Survey provide information on the number of bribes paid during the period covered, but no indication of how they are distributed over time. In addition, the time interval of surveys on corruption tends to be relatively short (often a year). Such data make it difficult to differentiate between corruption that appears randomly throughout departments and is eradicated and chronic corruption engrained in institutions. Yet the two types may require different policy responses. Subsequent work will extend the framework presented here to include measures and axioms that distinguish transient from chronic corruption.

REFERENCES


