



RESEARCH NOTE

# Methodological errors in corruption research: Recommendations for future research

Andrew Delios<sup>1</sup>,  
Edmund J. Malesky<sup>2</sup>, Shu Yu<sup>3,4</sup>  
and Griffin Riddler<sup>2</sup>

<sup>1</sup>NUS Business School, National University of Singapore, Singapore, Singapore; <sup>2</sup>Political Science Department, Duke University, Durham, NC, USA; <sup>3</sup>Suzhou Industrial Park Monash Research Institute of Science and Technology, Suzhou, China; <sup>4</sup>Department of Management, Monash University, Melbourne, Australia

**Correspondence:**

S Yu, Suzhou Industrial Park Monash Research Institute of Science and Technology, Suzhou, China  
e-mail: shu.yu@monash.edu

**Abstract**

The secretive, illegal, multidimensional, and ubiquitous nature of corruption leads to formidable difficulties in research design and measurement. When research fails to account for these challenges, it can lead to an empirical misalignment with concepts and theories of corruption, with inferential errors commensurately emerging. We define, measure, and track four common measurement errors and two common research design errors for papers on corruption published in international business/management and political economy journals in the 2000–2021 period. Our data marks a substantial opportunity to tighten the fit between theory and methods. We offer recommendations to accelerate improvements in empirical research on corruption, and indeed for other phenomena that are characterized by legal, moral, and social desirability concerns. These empirical recommendations contribute to more robust theory building.

*Journal of International Business Studies* (2024) 55, 235–251.  
<https://doi.org/10.1057/s41267-023-00637-8>

**Keywords:** corruption; ethics; bribery; measurement error; research design; list experiments; political economy

## INTRODUCTION

Corruption is the misuse of public power for private benefit. It often occurs where public and private sectors meet (Bardhan, 1997). Corruption is present in all geographies, and endemic in certain areas of the world, challenging citizens, managers, and public-sector officials alike. Unsurprisingly, corruption is a phenomenon of interest for scholars from many disciplines including international business (IB) and management, and political economy, which we define as research in political science and economics studying the relationship between the state and economic actors.

Even with substantial multi-disciplinary interest, the secretive, illegal, and opaque nature of corruption makes its measurement and modeling challenging, leading to persistent weaknesses in research designs and the measurement of corruption (Wathne & Stephenson, 2021). We define and document six research design weaknesses, which are either specific to corruption research, or

common to IB research in general. The first two are research design errors: (1) not accounting for endogeneity when modeling corruption, and (2) not guarding against respondents being socially predisposed to inaccurately report the incidence and scale of corruption. The next four are measurement errors: (3) aggregating conceptually distinct forms of corruption into a single measure, (4) not aligning the measure of corruption with the appropriate actor, (5) conflating perceptual and experience-based measures of corruption, and (6) misinterpreting the frequency of corruption as its scale.

Our evidence for the sustained prevalence of these six errors comes from our empirical review of research in the IB/management and the political economy literatures, published in the 2000–2021 period. Our review identifies how previous research in IB/management has heavily relied upon a small set of secondary source survey data with attendant potential problems for systematic measurement error. Aside from documenting longitudinal trends in the prevalence of these six weaknesses, we identify process-related issues and empirical tools that can enable researchers to reduce the prevalence of such errors, and thereby create a tighter alignment between theory and empirics.

Limitations exist in all research, but we contend that as the corruption literature currently stands, improved design and measurement are critical to aid in theoretical development of the novel relationships in which scholars are interested (Bahoo, Alon, & Paltrinieri, 2020; Cuervo-Cazurra, 2016). Moreover, heightened empirical acuity empowers scholars to undertake research that can enable firms to devise strategies to reduce the risks and costs of dealing with corruption, while facilitating states to develop anticorruption initiatives.

In our review, we first describe our process to identify and code research design challenges common to corruption studies. Next, we define and measure the six methodological errors in corruption research in the 2000–2021 period. We close by offering recommendations to improve empirical work.

## RESEARCH BACKGROUND

We focused our attention on research design, as the measurement and modeling of corruption are uniquely challenging compared to other IB phenomena, thus increasing the likelihood of errors emerging in methods and measurement (Wu,

Delios, Chen, & Wang, 2023). Corruption is challenging to measure for several reasons. First, corruption is both an antecedent and a consequent construct (Cuervo-Cazurra, 2016). It is nested in relationships with other institutional environment variables, such as voice and accountability, the rule of law, political risk, governance, and the degree of regulation among others (Treisman, 2007). As such, isolating the role of corruption as an independent variable, or defining it effectively as a dependent variable is challenging (Olken & Pande, 2012), heightening concerns about endogeneity and causal inference (Hill, Johnson, Greco, O'Boyle, & Walter, 2021; Li, Ding, Hu, & Wan, 2021; Reeb, Sakakibara, & Mahmood, 2012).

Next, corruption is generally considered to be illegal and immoral. Because of its illicit and illegal nature, people involved in a corrupt relationship are reluctant to admit it at all (Coutts & Jann, 2011), or to expose the full extent of corruption in a relationship (Gago-Rodríguez, Márquez-Illescas, & Núñez-Nickel, 2020). Hence, respondents typically do not report corruption, or report it incompletely or inaccurately (Jensen, Li, & Rahman, 2010), as answering honestly could place a respondent in danger of punitive actions (Jensen & Malesky, 2018).

A related complication is that corruption generally cuts across segments of societies. It can affect common citizens in day-to-day life, or businesspeople, such as managers, via small bribes for licenses, and it implicates elite leaders via grand corruption for major contracts. Hence, it is critical that measures of specific forms of corruption are tied directly to the correct respondents (Beaudry & Gile, 2020; Cycyota & Harrison, 2006). Safeguards are necessary to ensure that researchers connect to the actors most familiar with the specific corrupt behaviors about which they are theorizing.

Another unique feature of corruption is its opaqueness. As an example, diagnosing corruption from the auditing of financial reports of companies is complicated because companies hide such expenditures through generic reporting (e.g., entertainment expenditures) or omit it. As we cannot directly observe its incidence, respondents often infer the existence of corruption either indirectly by economic outcomes or by inaccurate attributions, instead of by direct experience.

Finally, corruption is multidimensional. Definitions of corruption disaggregate it into conceptual distinctions such as by size (petty versus grand) (Rose-Ackerman, 2013; Sartor & Beamish, 2018), by

whether it involves theft from the state (collusive versus non-collusive) (Foellmi & Oechslin, 2007; Guriev, 2004); by the recipient (politician versus bureaucrat) (Ang, 2020), by the victim (individual citizens, domestic firms, foreign investors), by the level of government in which it originates (centralized versus decentralized) (Fisman & Gatti, 2002), and by the specific, corrupt business–state interaction, such as facilitation, regulatory inspections, procurement (Ang, 2020; Rose-Ackerman, 2013; Shleifer & Vishny, 1993). Researchers need to be cognizant of the type of corruption being measured and if the indicated value connects to the frequency of the event or its scale, for example.

### Review Design

We conducted our review by comparing longitudinal trends in empirics for papers published in IB and management journals with longitudinal trends in empirics for papers published in the political economy literature.<sup>1</sup> We selected political economy for comparison as research in IB has a history of tight alignment with research from economics as well as from political science, making it a sensible adjacency.

We examined papers published in the 2000–2021 period in 25 journals. Thirteen journals were from IB/management, and are common to reviews in management (Aguilera, Marano, & Haxhi, 2019; Cuervo-Cazurra, Gaur, & Singh, 2019; Nippa & Reuer, 2019).<sup>2</sup> We added the *Journal of Business Ethics* given its focus on research on corruption. For the twelve leading journals in political economy, we included those used in recent methodological critiques (Card & DellaVigna, 2013; Hainmueller, Mummolo, & Xu, 2019; Heckman & Moktan, 2020; Montgomery, Nyhan, & Torres, 2018).<sup>3</sup> We searched these journals using the keywords “corruption” and “bribery,” as supplemented by a manual search of all identified articles. We identified 107 articles from the IB/management literature and 91 articles from the political economy literature.

### Review of Methods in Corruption Research

As reinforcement for our study’s motivation, Cuervo-Cazurra (2016: 37–38) noted problems with corruption measurement. Further, several secondary data sources<sup>4</sup> used with great frequency in corruption research yield different outcomes in country rankings by level of corruption. Judge, McNatt and Xu (2011) and Bahoo et al. (2020) noted limitations in corruption measures, but

empirical issues were not central to these reviews, nor were remedies offered.

Our first objective is to document trends in research methods by the six sources of error in research design or measurement. We start our assessment of research methods from concerns about causal inference (Aguinis, Ramani, & Cascio, 2020; Li et al., 2021), which is a concern that cuts across IB phenomena. We augment this assessment with social desirability concerns, which is a pernicious issue in corruption research. The social desirability bias that extends from corruption’s illegal and illicit nature makes respondents reluctant to respond truthfully to questionnaires or surveys about their actual experiences with corruption (Charron, 2016; Donchev & Ujhelyi, 2014; Olken, 2009).

Our next four items concern corruption measurement. Difficulties in gathering valid primary source data have led to a heavy reliance on secondary sources. Drawing on the unique characteristics of corruption, we identified four common errors in fit between measurement and theory: (1) aggregation error; (2) measurement inappropriate for actor types; (3) mismatch between perception and experience; and (4) mismatch between frequency and scale. We selected these four errors because our extensive literature review indicated that these errors were not only extremely common, but also had the most damaging impact on the inferences that can be drawn from corruption research. Our evidence quantifies the scope and nature of the two research design errors and four measurement errors in the methodological approaches employed in our survey of corruption studies.

### Research Design Errors

#### *Causal inference*

Our concern with causal inference follows an ever-increasing trend in scholarship that has sought to account for potentially crippling endogeneity errors between treatment and outcome (Li et al., 2021). There are a plethora of possibilities that make causal inference a prominent concern in corruption research. For example, does red tape lead to corruption or do bureaucrats generate red tape in order to engender corruption (Fredriksson, 2014; Guriev, 2004)? Research design issues are critical to disentangling conceptually related institutional dimensions, and to isolate the effects of corruption from the causes of corruption (Berliner & Erlich, 2015; Bussell, 2015; Cordero & Miller,

2019). For example, research on corruption must be able to isolate the effect of corruption from other institutional environment variables, such as ones that measure market attractiveness or governance (Brouthers, Gao, & Mcnicol, 2008). Critically, elements of the political environment, such as the tenure of a political party, can influence corruption levels, complicating the isolation of the effects of these two elements of the business environment (Cordero & Miller, 2019).

Fortunately for research on corruption, an abundance of tools exist to address concerns about causal inference and endogeneity. Authors should be aware of and use these tools to facilitate stronger causal inference. These tools include: (i) experimental methods (RCT, laboratory, or survey experiments), (ii) regression discontinuities, (iii) propensity score matching or entropy balancing techniques, (iv) difference-in-differences or generalized difference-in-differences using two-way fixed effects, and (v) instrumental variables analyses (Cunningham, 2021). Hill et al. (2021) provide a deep review of these and other tools. That said, herein, we reinforce these advocations for research with strong causal identification.

#### ***Social desirability error***

Self-reports of corruption can introduce error in measurement. Researchers need to take steps to ensure accurate reporting of bribes or other forms of corruption in their research design. Legal sanctions against corruption and cultural norms that discourage bribery will impel respondents to under-report or outright lie on the survey, leading corruption reports to be biased downwards. Researchers must recognize the errors that can emerge through self-reporting corruption and use cleverly designed primary data collection or try to find non-tainted data from secondary sources.

The challenge of social desirability error for corruption research is profound. For example, if we are concerned with understanding how trust in political institutions influences perceptions of corruption (Morris & Klesner, 2010), we would need to understand that the willingness to be candid to any queries about corruption, and avoid social desirability error, would extend in part from that same level of trust. Similarly, if we look at how international anti-bribery efforts, such as the US Foreign Corrupt Practices Act and the OECD Anti-Bribery Convention, may induce bribe reduction (Jensen & Malesky, 2018), extra-territoriality concerns from such agreements can lead to socially desirable but

false reports of reduced corruption, because admitting to such bribery can lead to punishment at home.

#### ***Aggregation error***

Aggregation error occurs when a measure is constructed from data collected from multiple respondents and often to multiple questions. Aggregate measures comprise several sub-types of corruption, which are conflated with each other. The consequence is that empirical tests can give ecological correlations, where coefficient estimates based on aggregate data lead to inflated estimates of lower-level relationships (Robinson, 1950).

Aggregation error hence occurs when researchers use an index measure to test a theory involving a specific corrupt behavior. Well-known secondary-source measures of corruption combine conceptually distinct forms of corruption, such as petty bribery during registration procedures, bribes during regulatory inspections, kick-backs on procurement contracts, or grand corruption. Often authors present a theory of one particular form of corruption but test it with an aggregate measure. For instance, a researcher may argue that one-stop-shop business registration reduces petty bribery when starting a business. However, this logic is tested by regressing an aggregate index of corruption on a measure of which countries have implemented one-stop-shops. If a negative correlation is observed, it is interpreted as evidentiary support, but we have no direct evidence that bribery dropped in business registration offices specifically. We only see a relationship between one-stop-shops and total corruption, which could have been caused by other institutional or policy changes occurring at the same time. Aggregation error raises the possibility of such spurious correlations in empirical analysis.

#### ***Measurement inappropriate for actor types***

Inappropriate types of actors refer to whether the units of observation for a specific measure map directly to the actors envisioned in the theory. If there is a mismatch between the actors in the theory and the units of observation, researchers will capture trends of corruption in one phenomenological setting, but with weakened confidence that is applicable to the researcher's theory.

For example, we will have an error representative of an inappropriate actor type if the theory being tested concerns the relationship between FDI and grand corruption, but the data collected are from



micro-level surveys concerning household bribery. If a paper presents a theoretical logic about a specific actor, but tests it on aggregate data that combines multiple actors (e.g., individual citizens, small and medium-sized enterprises, and large multinational firms), or tests it on a dataset where corruption responses are generated by different actors, we cannot know if the theorized actor is involved in the corrupt behavior at all.

#### ***Mismatch between perception and experience***

Respondents to survey questions that measure corruption often report on the basis of their perceptions of the situation, instead of their actual experiences with the situation. The pitfalls of such measurement are evident in the observation that country-level indicators of perceptions of bribery are correlated with levels of income, the quality of political institutions, and even religion, while reports of actual paid bribes are not (Olken, 2009; Treisman, 2007). If perceptions of corruption are correlated with income, for example, then an empirical test of the impact of a variable that is correlated with income on corruption will overestimate the strength of the relationship.

A considerable amount of research has been conducted with measures that are founded on perceptions of corruption. The error arises when theory concerns actual corruption experiences, but the test uses data measuring perceptions. This issue is important in the corruption literature because there is evidence that perceptions and experience do not align well.

Spencer and Gomez (2011: 299) measured an outcome of corruption using the BEEPs survey question, "How often *do firms like yours* nowadays need to make extra, unofficial payments to public officials for any of the following?" Although asking respondents to project away from themselves by answering about similarly situated firms is a common technique for reducing social desirability error, it introduces the concern that perceptions drive responses. To make the point clear that perceptions and experience do not align, 68% of firms in the 2009 BEEPs dataset answered the question about paying bribes in court (Ecaq41b), even though they had never actually been involved in a court case, either as a plaintiff or a defendant (Ecaq31e), and therefore possessed no actual experience with bribery in that setting.

#### ***Misinterpreting frequency of corruption as its scale***

Research has identified important differences between the incidence of corruption and the scale of corruption, such as in the distinction between petty and grand corruption (Rose-Ackerman, 2013; Rose-Ackerman & Palifka, 2016). We hence need to be acutely concerned about whether measurement choices fit with theory. If the researcher presents a theory of the size of corruption (i.e., firms not subject to the OECD anti-bribe convention pay more in bribes), we need to have the researcher test this theoretical idea using a measure of size (i.e., how much does a bribe cost?), not a measure of frequency (i.e., what share of firms engage in bribery?).

Empirical and conceptual imprecision in the frequency and scale distinction weakens confidence in whatever results are derived in the subsequent analysis. The conflation between frequency and scale is of critical prominence in some theoretical debates. For example, scholars classify different regimes by how predictable corruption is. These arguments rely on the fact that corruption is dissimilar to other government interventions, such as taxation, because of the illegality of the exchanges (Shleifer & Vishny, 1993; Wei, 1997). If a firm must pay a bribe to enter a market, and there is significant uncertainty that the necessary permits or licenses will be approved in response to the bribe payment, then firms will be less likely to enter that market. Given a particular scale of corruption, countries that have a more predictable system of corruption should see more investment, due to the reduction in uncertainty (Campos, Lien, & Pradhan, 1999). More predictable corruption, in turn, may also lead to reduced variability in the size of bribes. Because the variance in the amount of the bribe schedule is so important to theory, a measure of bribe frequency is inappropriate as a clean test of this theory.

### **ASSESSING THE PREVALENCE OF THE SIX ERRORS**

We evaluated the prevalence of the six research methods errors in our empirical review of the 198 studies. In terms of process, first, we met as a team to generate, discuss, and establish a shared understanding of the meaning of each error. Second, we developed a working definition for each of the six errors. Third, to establish a consistent implementation for the evaluation of each error, we completed a trial where each analyst independently

reviewed five manuscripts using the six criteria. The analysts compared assessments to cross-check each other’s coding process for alignment. Fourth, two analysts independently coded each paper. After independent review, the two analysts compared results. Discrepancies were resolved through a joint consultation with a third analyst. Fifth, as a random check, several papers were shared with another independent analyst to ensure that the coding was implemented consistently. Sixth, we archived details of our analysis for each paper.<sup>5</sup>

**Trend Analysis in the Six Errors**

The trends in errors in corruption research are based on 198 empirical manuscripts published in the 2000 to 2021 period. Figure 1 plots the distribution of article counts by 2-year intervals. The number of corruption-related articles in both IB/management and political economy journals increased steadily year-over-year.

Next, Figure 2 shows the mean number of errors by paper, year, and journal type. The number of errors decreased from a mean of four in the early 2000s to a mean of three, by the early 2010s. Further, research in political economy continued to show a decrease to a mean of one error per paper by the late 2010s. However, research in IB/management stayed at 2–3 errors per paper.

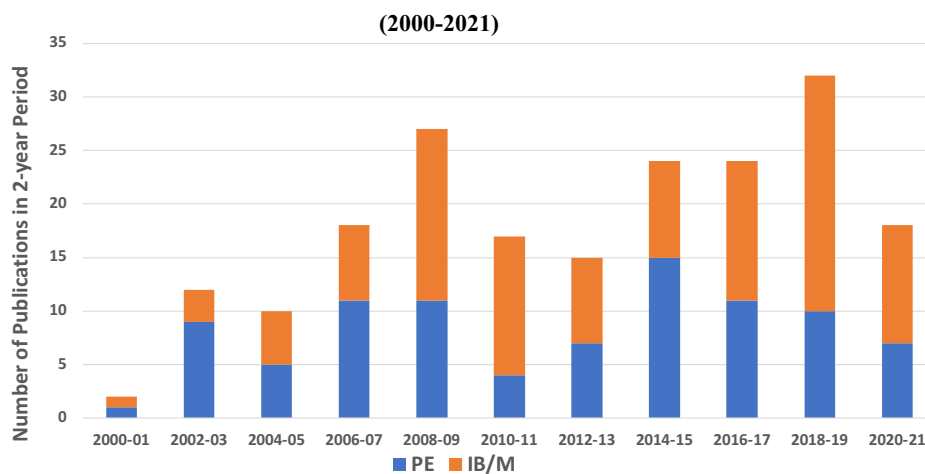
Next, we display the percentage of papers having a particular error by period and by field (Tables 1 and 2). In both IB/management and political economy publications, aggregation errors were encountered with similar incidence in 2000–2004. That said, both

fields showed improvements, with less than 15% of papers having an aggregation error in the 2015–2021 period. Types of actors’ error emerged in just 6% of papers in political economy and 24% of papers in IB/management by 2015–2021.

The prevalence of the perception versus experience error falls from 69% in the 2000–2004 period to 14% in the 2015–2021 period in political economy. In the same period for IB/management, the drop is from 71 to 34%. Meanwhile, the frequency versus scale error is sticky. This error was found in 100% of papers in IB/management and 77% of papers in political economy in the 2000–2004 period, but had only declined to about half of all papers in IB/management (46%) and political economy (57%) by the most recent period.

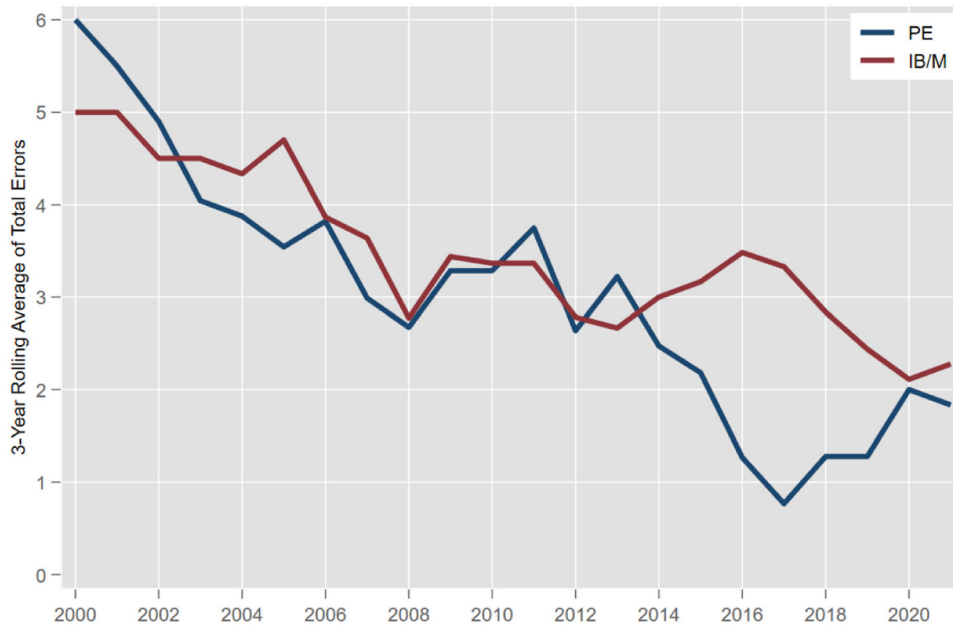
Next, causal inference and social desirability errors were between 71 and 92% for both political economy and IB/management in the 2000–2004 period. Causal inference error dropped to 23% and social desirability error dropped to 34% in the political economy sample by 2015–2021. However, in IB/management, causal inference error and social desirability error were at 70–72% in 2015–2021.

Finally, we compared trends by disaggregating comparisons by papers published in IB journals with those published in management, political economy journals. In sum, this three-category comparison identifies IB journals having a higher rate of incidence of each type of error than either management or political economy journals. As



PE refers to a publication in a Political Economy journal. IB/M refers to a publication in an International Business or Management journal.

**Figure 1** Temporal distribution of sample by journal type (2000–2021). Note PE refers to a publication in a political economy journal. IB/M refers to a publication in an international business or management journal.



**Figure 2** Mean score of the six errors by year and journal type. *Note* PE refers to a publication in a political economy journal. IB/M refers to publications in an international business or management journal.

**Table 1** Prevalence of errors in corruption research (international business/management)

Type of error	Percentage of papers having errors in period			
	2000–2004	2005–2010	2011–2015	2015–2021
1. Causal inference	86	92	72	72
2. Social desirability	71	84	68	70
3. Aggregation	43	40	36	12
4. Types of actors	86	48	44	24
5. Perception and experience	71	28	68	34
6. Frequency and scale	100	64	56	46

**Table 2** Prevalence of errors in corruption research (political economy)

Type of error	Percentage of papers having errors in period			
	2000–2004 (%)	2005–2010 (%)	2011–2015 (%)	2015–2021 (%)
1. Causal inference	92	50	32	23
2. Social desirability	85	67	74	34
3. Aggregation	69	50	58	14
4. Types of actors	46	25	42	6
5. Perception and experience	69	42	68	14
6. Frequency and scale	77	79	74	57

examples, in all three categories, aggregation errors were common in 2000–2004, but by 2015–2021, both management and political economy registered 15% of journals with an aggregation error

compared to a 25% prevalence in IB journals. If we look at the types of actors’ error, in 2015–2021, it emerged in 3% of management papers and 6% of political economy articles, but 58% of IB papers.

As for the prevalence of the perception versus experience error, it appeared in 80% of IB papers published in the 2015–2021 period, but in 14% of political economy journals and 6% of management journals. Then, as might be expected, the frequency versus scale error is sticky. This error declined to 80% of IB papers in 2015–2021, 57% of political economy papers, and 25% of management articles.

Lastly, causal inference and social desirability errors were high for all three disciplines in the 2000–2004 period. However, causal inference error dropped to 23% and social desirability error dropped to 34% in the political economy sample by 2015–2021. In management, causal inference error and social desirability error remained in two-thirds of publications, while in IB the incidence was 79% of published causal inference design errors and 89% missing shielding for social desirability concerns.

## RECOMMENDATIONS

Our focus is to understand the methods that typify corruption research to identify areas that can be improved and made better-suited to test theoretical innovations with new empirical work. We identify six errors. Although some errors, such as causal inference and aggregation, are common concerns in empirical research in IB, others such as social desirability bias, perception versus experience, frequency versus scale and actor type, are notably pronounced in corruption research. Moreover, some of these errors might also emerge in other areas of research that share characteristics similar to corruption in terms of illegality, moral acceptance, or social acceptance pressures. We elaborate on this point at the end of this discussion. Here, we first present an overview of the implications of our findings before moving to process-related and tool-related recommendations to improve empirical work on corruption.

### Tighten Focus on Reducing Research Methods Errors

Our analysis of the errors shows a decreasing prevalence of errors, which is a positive trend. That said, while also acknowledging that there will be limitations to any piece of empirical research, we note that across the 85 papers that were published between 2015 and 2021, we find an interquartile range of 1–3 errors per paper, with 34% of papers still having three or more errors. Moreover, the rate of improvement in IB/M has been less rapid than

that in the political economy literature. Forty percent of political economy papers published in the 2015–2021 period suffer from more than one error, but that figure climbs to 66% in the IB/management sample in the same period. van Witteloostuijn and van Hugten (2022) similarly found economics-oriented studies to lead in the implementation of methodological improvements when compared to other social science disciplines.

Finally, although we identify several errors that we expect to be more common to research on corruption than to research in IB in general, even common issues, such as addressing casual inference, remain disturbingly high at 72% of all papers published in the 2011–2021 period. This stickiness of causal inference-related issues occurs despite exhortations in JIBS (Li et al., 2021; Reeb et al., 2012) and other journals (Hill et al., 2021) to not only address it, but also provide clear advice on how to address endogeneity. There are clear opportunities to improve empirical research on corruption.

### Process-Related Recommendations for Improving Research Designs

Authors, reviewers, and editors need to push corruption research to reduce the prevalence of errors, which ideally is an issue that is addressed at the design stage of the research; not post-data collection. Given the resource-related and validity-related challenges to collecting good primary source data, it is understandable that researchers rely extensively on common archival sources of cross-national panel data on corruption. We acknowledge the research exigencies that can drive such choices, but we encourage scholars to follow a research process that can help to reduce method errors from a research design perspective. To that end, we offer Table 3, which identifies six process-related steps that help identify and reduce the prevalence of each type of error we identify.

Although process-related issues to research design might seem to be self-evident, our data indicates that reminders are needed to guide researchers through not only ever-present challenges to effective research design such as addressing causal ambiguity, but also to challenges that are more pronounced in corruption research than in other IB research domains. As such, scholars investigating corruption need to be uniquely attuned to steps 2 through 6, which provide checks in the processes related to countering social desirability concerns (step 2), neutralizing aggregation



**Table 3** Process-related recommendations to reduce research methods errors

Process	Practices to avoid	Practices to encourage
<i>Step 1</i> If using observable data, does it use a causal research method?	Interpreting econometric models as causal fact. Arguing that including control variables accounts for problems of endogeneity. Articulating vague concerns about endogeneity without correctly specifying potential pathways of bias.	Utilize experimental research methods where applicable. OR Utilize one of the following designs: Instrumental variables Difference-in-differences Regression discontinuity design Simultaneous equation models Matching
<i>Step 2</i> If using survey data, does it properly shield respondents to avoid social desirability bias?	Relying on direct or indirect questions that ask about sensitive topics without shielding	Use confidentiality envelopes or boxes to separate sensitive questions. Develop shielded response techniques, such as list experiments.
<i>Step 3</i> If using a nuanced conception of corruption, does the measure align with the appropriate level of aggregation?	Using country-level indices to test micro-level theories of corruption. Using composite measures that combine different aspects of corruption (petty vs. grand, collusive vs. coercive, etc.).	Explicitly map the specifics of corruption onto the chosen measure. Disaggregate existing indices. Draw on more specified data from administrative sources or official filings.
<i>Step 4</i> If the theory of corruption concerns a specific actor type, does the respondent for the measure align? <i>Step 5</i> Does the theory focus on actual levels of corruption (experience) or perceptions of corruption?	Using country-level indices to test theories specific to certain forms of corruption (household bribery, firm corruption, etc.). Using perception-based surveys to measure levels of corruption. Using indirect questions to measure a firm or individual's experience of corruption.	Draw on more specified data from administrative sources, firm surveys, or other novel resources. Use direct observation to measure corruption. Draw on secondary source data (customs reports, official disclosures, etc.). When dealing with perceptions of corruption, consider utilizing lab or survey experiments.
<i>Step 6</i> Does the theory focus on how often corruption occurs, or does it involve the amount of money involved?	Conflating the frequency of bribe payments with the scale of payments.	When measuring scale, include questions that directly deal with money involved in corrupt transactions. Draw on novel methods, such as list experiments.

concerns (step 3), identifying respondents who match the underlying theory being tested (step 4), addressing perceptions versus experience concerns (step 5), and being sure to match measures and conceptualizations of corruption along the frequency versus scale criterion (step 6).

For each of these steps, we articulate not only practices to avoid, but also advocate specific positive practices to implement. For example, for step 5, where we are concerned with experience versus perception, we clearly indicate that the phrasing of surveys should be aimed away from perception-based indicators and should not involve the use of indirect questions. Instead, the research should be

designed to observe incidences of corruption directly, or cleverly use secondary source data, or as a third option, utilize lab or survey experiments. That said, we go a step further in our recommendations to identify tools that can help to implement these process-related recommendations.

### Technique-Related Recommendations for Improving Research Designs

Addressing the research methods errors we identify requires a tight attention to research design and a need for scholars to further trade-off expediency for research rigor. To facilitate the operationalization of this exhortation, we present Table 4. This table identifies nine techniques for improving the

execution of research on corruption. The techniques range from primary to secondary sources, and also cover the category of shielded source techniques, which primarily deal with survey data. Hence, options exist for scholars who wish to operate with archival or self-sourced data.

As we are cognizant of the trade-off between expediency and resources and research effectiveness, we illustrate the general requirements and limitations of each technique, as compared to the specific error(s) it helps to remedy. We note that this table is specific to measures and issues that are more pervasive in corruption research than in other fields of phenomenological endeavor in IB research. Notably, the right-most column provides references to studies that have implemented the relevant technique to enable scholars to quickly reference and potentially implement the same technique.

When we combine the information in Tables 3 and 4, we have developed a process for identifying potential sources of error, and then techniques to address those sources of error. For example, if we believe social desirability bias will be a major concern, which is common in corruption research, then we return to the concern about whether respondents are properly shielded. Fortunately, as shown in Table 4, solutions to countering social desirability concerns exist for both primary and secondary source studies: (i) the creative use of administrative data (Jeong & Siegel, 2018), (ii) direct observation (McMillan & Zoido, 2004; Olken & Barron, 2009), and (iii) shielded survey techniques (SSTs) to guarantee anonymity to respondents (Coutts & Jann, 2011). For example, as in Table 4, scholars in IB/management have executed studies that creatively use administrative data (Jeong & Siegel, 2018).

An example of the creative use of secondary sources exists in how variances in tax or customs reports can highlight illicit money flows, because discrepancies in official disclosures can show that firms must pass off the cost of bribes to move goods across borders (Fisman & Wei, 2004; Hough, 2017). Meanwhile, for primary source measures, researchers could operate with (i) confidentiality envelopes or boxes that separate sensitive questions from the survey, and (ii) shielded response and list experiments, which can reduce social desirability error when surveys have questions about bribe experiences (Corbacho, Gingerich, Oliveros, & Ruiz-Vega, 2016; Coutts & Jann, 2011; Malesky, Gueorguiev, & Jensen, 2015).

To continue down this path, we highlight how shielded response techniques are becoming more common in the social sciences, including public health (Starosta & Earleywine, 2014) and education (Walsh & Braithwaite, 2008). A shielded response enables primary data to be gathered while minimizing social desirability concerns. A common format is the unmatched count technique (UCT) or list experiment (Ahart & Sackett, 2004; Coutts & Jann, 2011; Gonzalez-Ocantos, De Jonge, Meléndez, Osorio & Nickerson, 2012).

In a list experiment, respondents are asked whether they engaged in a number of different activities. If bribery is of concern for the researcher, then bribery will be only one of several activities provided in the description. As respondents are asked how many of the listed activities they have completed, there is plausible deniability about their engagement in corruption, as they have not specified directly that corruption was one of the activities in which they have engaged.

Importantly, a list experiment allows for the detection of corruption *frequency*, but it has typically not been utilized to measure the *scale* of corruption, which raises the specter of the frequency versus scale error emerging. But as our last row in Table 4 indicates, list experiments can be adapted to measure scale. Appendix 1 provides an example of how to implement a list experiment to measure scale. In short, to measure the scale of corruption, respondents are asked to report on the total cost of a group of listed activities. The control group receives a list that includes a rare but plausible activity that is not corruption, and the treatment group receives the same list, except that the rare but plausible activity is replaced by corruption. Comparison of the treatment and control groups allows for the calculation of the average price of a bribe, which can also help untangle theoretical concerns about grand and petty corruption (Malesky, Dulay & Peltovuori, 2020). Hence, list experiments can measure the frequency and scale of corruption, in addition to accounting for social desirability and causal identification.

### Extensions to Our Recommendations

Our recommendations focus on solutions to research design errors common to primary and secondary source research on corruption. At the same time, the errors we identify can be considered as common to research in IB in general, for example causal ambiguity, or of heightened prominence to corruption research, for example social desirability

**Table 4** Technique-related recommendations to reduce research methods errors

Technique	Description	Requirements & limitations	Errors remedied	Selected examples
<i>Primary source techniques</i>				
Direct enumeration	Enumerators are trained and embedded with selected populations to observe and report on corrupt transactions.	Requires a substantial investment of resources and manpower.	Aggregation error Inappropriate measurement	McMillan and Zoido (2004); Olken and Barron (2009)
Direct interviews	The researchers directly interview stakeholders and those affected by corruption, including managers, workers, officials, and citizens.	May be difficult to obtain honest answers from interviewees. Difficult to yield data for strictly quantitative methods.	Aggregation error Inappropriate measurement	Collins, Uhlenbruck, and Rodriguez (2009) Gans-Morse (2017) Tellez, Wibbels, and Krishna (2020)
Field experiments	Researchers develop a proposed treatment and implement it either in the policy space or by altering the behavior of citizens or firms. Researchers then directly enumerate the effects on corruption and other types of behaviors.	Requires co-operation with government officials and/or citizens.	Causal inference Aggregation error Inappropriate measurement	Olken (2007) Robinson and Seim (2018)
Survey experiments	Within a survey, different respondents are randomly assigned to experimental conditions/treatments. These treatments can be story vignettes or prompts that then trigger respondents for further questions.	Often involves the use of hypotheticals, potentially limiting real-world applicability.	Causal inference Inappropriate measurement Aggregation error	Huang (2015) Boas, Hidalgo, and Melo (2019)
<i>Secondary source techniques</i>				
Administrative data	Researchers draw on published administrative data, such as government contracts or administrative budgets.	In many contexts, these data are not publicly available.	Aggregation error Inappropriate measurement Frequency vs. scale	Fisman and Wei (2004) Gorodnichenko and Peter (2007) Jeong and Siegel (2018)
Official disclosures and filings	This research utilizes purchased or publicly available data on firms or politicians, often with a reliance on well-theorized proxies for corruption.	The content and accuracy of filings are dependent on the laws and norms of the locale.	Aggregation error Inappropriate measurement	Sampath, Gardberg, and Rahman (2018) Sari, Cahaya, and Joseph (2021) Fisman, Schulz, and Vig (2014)
<i>Shielded survey techniques</i>				
Confidentiality envelopes	Respondents fill out the "sensitive" portion of the survey on a separate ballot and seal it in an envelope.	High logistical burden to implement Requires trust of respondent	Social desirability	Kotz (2004)
List experiments (frequency type)	Respondents are asked whether or not they engaged in a number of activities. Both groups receive a list of the same number of activities: the control group has a rare but plausible activity listed, while that activity is replaced by corruption for the treatment group.	Respondent must want to reveal sensitive behavior Administration requires multiple versions of survey instruments	Aggregation error Social desirability	Gonzalez-Ocantos et al. (2012) Corbacho et al. (2016) Jensen and Malesky (2018)

Table 4 (Continued)

Technique	Description	Requirements & limitations	Errors remedied	Selected examples
List experiments (scale type)	In addition to questions about the frequency of corruption, respondents are asked to report on the total cost of a group of listed activities. The same difference method is utilized to compare the control and treatment groups.	Respondent must have an accurate accounting of costs. Requires strong knowledge of common activities and their costs.	Social desirability Inappropriate measurement Frequency vs. scale	Malesky et al. (2020)

and experience versus perceptions. Yet, it is also possible that these same errors that are common in research in corruption can be found in other areas. For example, regulatory violations can occur in domestic and international settings. Research on a firm's lack of compliance with labor safety, consumer safety, or environmental standards can be subject to social desirability concerns, frequency versus scale issues, and aggregation errors. As such, techniques such as shielded responses can be helpful. Along a similar vein, research on corporate social responsibility (CSR) is a domestic and international issue (Napier, Knight, Luo, & Delios, 2023). Research on CSR ratings highlights a tendency of firms to misrepresent (exaggerate or understate) social performance (Kim & Lyon, 2015; Wang, Jia, & Zhang, 2021). Creatively using administrative data to compare multiple CSR reports can help to overcome errors found in the aggregation of CSR data.

Next, we note here that our empirical review captured quantitative methods. The reason for this is not a reflection of epistemological preferences or biases, instead it emerged directly from the 198 studies, which were overwhelmingly quantitative in design. That said, the same concerns about social desirability, actor type, scale and frequency and aggregation could emerge in qualitative research.

We fully expect that qualitative research on corruption can yield substantive insight into the phenomenon, but at the same time, social desirability and researcher access challenges are profound. That said, qualitative research can be pursued as part of a mixed-methods design as in Ang (2020), who proposed a two-dimensional definition of corruption: whether it involves theft from public coffers or exchanges between firms and officials, and whether it operates at ground-level bureaucrats or the political elite level. Such qualitative research can provide substantial insight into organizational practices in corrupt settings (Hoang,

2018) or on the nature of tight relationships between business and political elite (Apriliyanti, Dieleman & Randøy, 2023).

There are two tangencies to the above discussion. First, aside from the political economy literature, researchers can seek to draw inspiration for research design to counter errors from other areas of social science research. An example from psychology shows how experiments can help counter errors. De Waele, Weißmüller and van Witteloostuijn (2021) conducted a quasi-experiment to understand the causal mechanisms of bribery behavior. Causal inference was enabled by their experimental approach, which permitted strong inference to be made about how corruption is influenced by context.

Our second tangency connects to the idea that these examples and others we cite in this paper are not directly related to IB research. This observation can raise the question about whether corruption and the ideas we raise herein are relevant to IB scholars. For the first issue, unquestionably, corruption is of theoretic and tangible interest to IB scholars. The phenomenon is a domestic one and an international one, but the secretive and illicit nature of corruption makes it not only a challenging one for IB practitioners to manage, but it also makes it a formidable one for IB scholars to study. For example, the issue of social desirability will be of greater prominence when an international element is involved, given the greater barriers to the creation of trust and the pronounced challenges to building local resources and networks to study corruption effectively. As a result, learning from best practices in adjacent disciplines and from domestically oriented studies can help reduce the resource- and time-related hurdles to conducting research on corruption in IB settings.

Indeed, IB scholars are uniquely positioned compared to other social scientists to dive deep into the connections between business environments and

corruption to understand firms' strategic behaviors with respect to corruption. Enhanced precision in the conceptualization and measurement of corruption is critical to seize this opportunity. By blending the questions that IB scholars are well-positioned to ask with the methodological innovations that are progressing across the social sciences, IB scholars have a greater opportunity to not only conduct more rigorous studies but to also have IB contribute to other disciplines (Aguinis & Gabriel, 2022; Cantwell, Piepenbrink, & Shukla, 2014).

Relatedly, with progress, IB research can help leaders manage issues filled with ethical conundrums and legal ambiguities, as exemplified by corruption. Hence, if we understand the phenomenon better, managers can address corruption through their own agency. Our focus on errors aligns with our recommendation that corruption research should build an empirical foundation that aids academics in their efforts to develop knowledge about what is effective for managing specific types of corruption, and how some of the moral and legal knots can be untangled to proactively contend with corruption.

### CONCLUSION

Our review yields a series of actionable points about how empirical research on corruption can advance, to better enable researchers to test emergent theories of corruption. More so than launching a methodological critique, we conducted our empirical review to heighten the awareness of scholars about innovative research designs for theory testing in the corruption literature. We believe it is critical for the research methods deployed in corruption studies to aim to be largely free of the six errors we have identified in order to effectively test the nuances in emergent theories and ideas on corruption. We outline process-related steps that can reduce the prevalence of these errors at the research design stage. We follow this with an introduction of techniques drawn from other research disciplines that can enhance the rigor of IB-oriented research on corruption, while carefully considering trade-offs between time and resources and the reduction of errors. Accordingly, we recommend that researchers (1) be explicit in countering potential research design and measurement errors, (2) continue to invest in new sources of primary and secondary data, and (3) engage with the forefront to bring innovations in empirical methods into their own research.

### ACKNOWLEDGEMENTS

We thank Pasha Mahmood, Kulwant Singh, Bo Nielsen and Markus Taussig for discussions that stimulated some of the ideas and views in this submission. We are also grateful for the candid advice from three reviewers and our editor to help us progress this work in numerous important ways. The earliest version of this review paper benefited greatly from the comments of participants at the November 19–21, 2018, IB Workshop hosted by the Discipline of International Business, University of Sydney. We are thankful for the excellent research assistance of Jimmy Z. Chen and Tao Wu. All ideas and errors in this submission remain the responsibility of the authors.

### NOTES

<sup>1</sup><https://www.palgrave.com/gp/journal/41267/authors/call-for-papers-and-proposals>

<sup>2</sup>Academy of Management Journal, Academy of Management Review, Global Strategy Journal, International Business Review, Journal of Business Ethics, Journal of International Business Studies, Journal of International Management, Journal of Management, Journal of Management Studies, Journal of World Business, Management International Review, Organization Science, Strategic Management Journal.

<sup>3</sup>American Economic Review, American Journal of Political Science, American Political Science Review, Comparative, Political Studies, Econometrica, International Organization, Journal of Politics, Journal of Political Economy, Quarterly Journal of Economics, Quarterly Journal of Political Science, Review of Economic Studies, World Politics.

<sup>4</sup>Corruption Perceptions Index (CPI) (<https://www.transparency.org/en/cpi/2020/index/nzl>), World Bank's World Business Environment Survey, Business Environment and Enterprise Performance Survey (BEEPS) (<https://www.beeps-ebd.com/data/>) World Bank's Worldwide Governance Indicators (WGI) database (<https://info.worldbank.org/governance/wgi/>), and United Nations Conference on Trade and Development (UNCTAD) (<https://unctadstat.unctad.org/EN/>).

<sup>5</sup>The table can be found here: <https://dataverse.harvard.edu/dataverse/emalesky;jsessionid=c2b8d8112b819d26abde2a8cd22c>



## REFERENCES

- Aguilera, R. V., Marano, V., & Haxhi, I. 2019. International corporate governance: A review and opportunities for future research. *Journal of International Business Studies*, 50: 1–42.
- Aguinis, H., & Gabriel, K. P. 2022. International business studies: Are we really so uniquely complex? *Journal of International Business Studies*, 53(9): 2023–2036.
- Aguinis, H., Ramani, R. S., & Cascio, W. F. 2020. Methodological practices in international business research: An after-action review of challenges and solutions. *Journal of International Business Studies*, 51(9): 1593–1608.
- Ahart, A. M., & Sackett, P. R. 2004. A new method of examining relationships between individual difference measures and sensitive behavior criteria: Evaluating the unmatched count technique. *Organization Research Methods*, 7(1): 101–114.
- Ang, Y. Y. 2020. *The paradox of economic boom and vast corruption*. Cambridge: Cambridge University Press.
- Apriliyanti, I. D., Dieleman, M., & Randøy, T. 2023. Multiple-principal demands and CEO compliance in emerging market state-owned enterprises *Journal of Management Studies*, Forthcoming.
- Bahoo, S., Alon, I., & Paltrinieri, A. 2020. Corruption in international business: A review and research agenda. *International Business Review*, 29(4): 101660.
- Bardhan, P. 1997. Corruption and development: A review of issues. *Journal of Economic Literature*, 35(3): 1320–1346.
- Beaudry, I. S., & Gile, K. J. 2020. Correcting for differential recruitment in respondent-driven sampling data using ego-network information. *Electronic Journal of Statistics*, 14(2): 2678–2713.
- Berliner, D., & Erlich, A. 2015. Competing for transparency: Political competition and institutional reform in Mexican states. *American Political Science Review*, 109(1): 110–128.
- Boas, T. C., Hidalgo, F. D., & Melo, M. A. 2019. Norms versus action: Why voters fail to sanction malfeasance in Brazil. *American Journal of Political Science*, 63(2): 385–400.
- Brouthers, L. E., Gao, Y., & Mcnicol, J. P. 2008. Corruption and market attractiveness influences on different types of FDI. *Strategic Management Journal*, 29: 673–680.
- Bussell, J. 2015. Typologies of corruption: A pragmatic approach. In S. Rose-Ackerman, & P. Lagunes (Eds.), *Greed, corruption, and the modern state* Cheltenham: Edward Elgar Publishing.
- Campos, J. E., Lien, D., & Pradhan, S. 1999. The impact of corruption on investment: Predictability matters. *World Development*, 27(6): 1059–1067.
- Cantwell, J., Piepenbrink, A., & Shukla, P. 2014. Assessing the impact of JIBS as an interdisciplinary journal: A network approach. *Journal of International Business Studies*, 45(7): 787–799.
- Card, D., & DellaVigna, S. 2013. Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1): 144–161.
- Charron, N. 2016. Do corruption measures have a perception problem? Assessing the relationship between experiences and perceptions of corruption among citizens and experts. *European Political Science Review*, 8(1): 147–171.
- Collins, J. D., Uhlenbruck, K., & Rodriguez, P. 2009. Why firms engage in corruption: A top management perspective. *Journal of Business Ethics*, 87(1): 89–108.
- Corbacho, A., Gingerich, D. W., Oliveros, V., & Ruiz-Vega, M. 2016. Corruption as a self-fulfilling prophecy: Evidence from a survey experiment in Costa Rica. *American Journal of Political Science*, 60(4): 1077–1092.
- Cordero, A. M., & Miller, S. R. 2019. Political party tenure and MNE location choices. *Journal of International Business Studies*, 50(6): 973–997.
- Coutts, E., & Jann, B. 2011. Sensitive questions in online surveys: Experimental results for the randomized response technique (RRT) and the unmatched count technique (UCT). *Sociological Methods Research*, 40(1): 169–193.
- Cuervo-Cazurra, A. 2016. Corruption in international business. *Journal of World Business*, 51(1): 35–49.
- Cuervo-Cazurra, A., Gaur, A., & Singh, D. 2019. Pro-market institutions and global strategy: The pendulum of pro-market reforms and reversals. *Journal of International Business Studies*, 50(4): 598–632.
- Cunningham, S. 2021. *Causal inference: The mixtape*. New Haven: Yale University Press.
- Cycyota, C. S., & Harrison, D. A. 2006. What (not) to expect when surveying executives: A meta-analysis of top manager response rates and techniques over time. *Organizational Research Methods*, 9(2): 133–160.
- De Waele, L., Weißmüller, K. S., & van Witteloostuijn, A. 2021. Bribery and the role and social value orientation: A multi-site experimental study in Belgium, Germany, and the Netherlands. *Frontiers in Psychology*, 12: 655964.
- Donchev, D., & Ujhelyi, G. 2014. What do corruption indices measure? *Economics & Politics*, 26(2): 309–331.
- Fisman, R., & Gatti, R. 2002. Decentralization and corruption: Evidence across countries. *Journal of Public Economics*, 83(3): 325–345.
- Fisman, R., Schulz, F., & Vig, V. 2014. The private returns to public office. *Journal of Political Economy*, 122(4): 806–862.
- Fisman, R., & Wei, S.-J. 2004. Tax rates and tax evasion: Evidence from “missing imports” in China. *Journal of Political Economy*, 112(2): 471–496.
- Foellmi, R., & Oechslin, M. 2007. Who gains from non-collusive corruption? *Journal of Development Economics*, 82(1): 95–119.
- Fredriksson, A. 2014. Bureaucracy intermediaries, corruption and red tape. *Journal of Development Economics*, 108: 256–273.
- Gago-Rodríguez, S., Márquez-Illescas, G., & Núñez-Nickel, M. 2020. Denial of corruption: Voluntary disclosure of bribery information. *Journal of Business Ethics*, 162(3): 609–626.
- Gans-Morse, J. 2017. Demand for law and the security of property rights: The case of post-soviet Russia. *American Political Science Review*, 111(2): 338–359.
- Gonzalez-Ocantos, E., De Jonge, C. K., Meléndez, C., Osorio, J., & Nickerson, D. W. 2012. Vote buying and social desirability bias: Experimental evidence from Nicaragua. *American Journal of Political Science*, 56(1): 202–217.
- Gorodnichenko, Y., & Peter, K. S. 2007. Public sector pay and corruption: Measuring bribery from micro data. *Journal of Public Economics*, 91(5–6): 963–991.
- Guriev, S. 2004. Red tape and corruption. *Journal of Development Economics*, 73(2): 489–504.
- Hainmueller, J., Mummolo, J., & Xu, Y. 2019. How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. *Political Analysis*, 27(2): 163–192.
- Heckman, J. J., & Moktan, S. 2020. Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, 58(2): 419–470.
- Hill, A. D., Johnson, S. G., Greco, L. M., O’Boyle, E. H., & Walter, S. L. 2021. Endogeneity: A review and agenda for the methodology-practice divide affecting micro and macro research. *Journal of Management*, 47(1): 105–143.
- Hoang, K. K. 2018. Risky investments: How local and foreign investors finesse corruption-rife emerging markets. *American Sociological Review*, 83(4): 657–685.
- Hough, D. 2017. *There’s more to measuring corruption than transparency international’s annual index, just released*. Available at <https://www.washingtonpost.com/news/monkey-cage/wp/2017/01/26/theres-more-to-measuring-corruption-than-transparency-internationals-annual-index-just-released/>.



- Huang, H. 2015. International knowledge and domestic evaluations in a changing society: The case of China. *American Political Science Review*, 109(3): 613–634.
- Jensen, N. M., Li, Q., & Rahman, A. 2010. Understanding corruption and firm responses in cross-national firm-level surveys. *Journal of International Business Studies*, 41(9): 1481–1504.
- Jensen, N. M., & Malesky, E. J. 2018. Nonstate actors and compliance with international agreements: An empirical analysis of the OECD anti-bribery convention. *International Organization*, 72(1): 33–69.
- Jeong, Y., & Siegel, J. I. 2018. Threat of falling high status and corporate bribery: Evidence from the revealed accounting records of two South Korean presidents. *Strategic Management Journal*, 39(4): 1083–1111.
- Judge, W. Q., McNatt, D. B., & Xu, W. 2011. The antecedents and effects of national corruption: A meta-analysis. *Journal of World Business*, 46(1): 93–103.
- Kim, E.-H., & Lyon, T. P. 2015. Greenwash vs. Brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, 26(3): 705–723.
- Kotz, S. 2004. Sealed envelope technique. *Encyclopedia of Statistical Sciences*.
- Li, J., Ding, H., Hu, Y., & Wan, G. 2021. Dealing with dynamic endogeneity in international business research. *Journal of International Business Studies*, 52(3): 339–362.
- Malesky, E., Dulay, D., & Peltovuori, V. 2020. *The Myanmar business environment index 2020: Measuring economic governance for private sector development*. The Asian Foundation. <http://hdl.handle.net/11540/12790>.
- Malesky, E. J., Gueorguiev, D. D., & Jensen, N. M. 2015. Monopoly money: Foreign investment and bribery in Vietnam, a survey experiment. *American Journal of Political Science*, 59(2): 419–439.
- McMillan, J., & Zoido, P. 2004. How to subvert democracy: Montesinos in Peru. *Journal of Economic Perspectives*, 18(4): 69–92.
- Montgomery, J. M., Nyhan, B., & Torres, M. 2018. How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*, 62(3): 760–775.
- Morris, S. D., & Klesner, J. L. 2010. Corruption and trust: Theoretical considerations and evidence from Mexico. *Comparative Political Studies*, 43(10): 1258–1285.
- Napier, E., Knight, G., Luo, Y., & Delios, A. 2023. Corporate social performance in international business. *Journal of International Business Studies*, 54(1): 61–77.
- Nippa, M., & Reuer, J. 2019. On the future of international joint venture research. *Journal of International Business Studies*, 50(4): 555–597.
- Olken, B. A. 2007. Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Public 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., & Barron, P. 2009. The simple economics of extortion: Evidence from trucking in Aceh. *Journal of Political Economy*, 117(3): 417–452.
- Olken, B. A., & Pande, R. 2012. Corruption in developing countries. *Annual Review of Economics*, 4(1): 479–509.
- Reeb, D., Sakakibara, M., & Mahmood, I. P. 2012. From the editors: Endogeneity in international business research. *Journal of International Business Studies*, 43(3): 211–218.
- Robinson, A. L., & Seim, B. 2018. Who is targeted in corruption? Disentangling the effects of wealth and power on exposure to bribery. *Quarterly Journal of Political Science*, 13(3): 313–331.
- Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3): 351–357.
- Rose-Ackerman, S. 2013. *Corruption: A study in political economy*. New York: Academic Press.
- Rose-Ackerman, S., & Palifka, B. J. 2016. *Corruption and government: Causes, consequences, and reform*. Cambridge: Cambridge University Press.
- Sampath, V. S., Gardberg, N. A., & Rahman, N. 2018. Corporate reputation's invisible hand: Bribery, rational choice, and market penalties. *Journal of Business Ethics*, 151(3): 743–760.
- Sari, T. K., Cahaya, F. R., & Joseph, C. 2021. Coercive pressures and anti-corruption reporting: The case of ASEAN countries. *Journal of Business Ethics*, 171: 475–511.
- Sartor, M. A., & Beamish, P. W. 2018. Host market government corruption and the equity-based foreign entry strategies of multinational enterprises. *Journal of International Business Studies*, 49(3): 346–370.
- Shleifer, A., & Vishny, R. W. 1993. Corruption. *The Quarterly Journal of Economics*, 108(3): 599–617.
- Spencer, J., & Gomez, C. 2011. MNEs and corruption: The impact of national institutions and subsidiary strategy. *Strategic Management Journal*, 32: 280–300.
- Starosta, A. J., & Earleywine, M. 2014. Assessing base rates of sexual behavior using the unmatched count technique. *Health Psychology and Behavioral Medicine*, 2(1): 198–210.
- Tellez, J. F., Wibbels, E., & Krishna, A. 2020. Local order, policing, and bribes: Evidence from India. *World Politics*, 72(3): 377–410.
- 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.
- van Witteloostuijn, A., & van Hugten, J. 2022. The state of the art of hypothesis testing in the social sciences. *Social Sciences & Humanities Open*, 6(1): 100314.
- Walsh, J. A., & Braithwaite, J. 2008. Self-reported alcohol consumption and sexual behavior in males and females: Using the unmatched-count technique to examine reporting practices of socially sensitive subjects in a sample of university students. *Journal of Alcohol and Drug Education*, 52(2): 49–72.
- Wang, H., Jia, M., & Zhang, Z. 2021. Good deeds done in silence: Stakeholder management and quiet giving by Chinese firms. *Organization Science*, 32(3): 649–674.
- Wathne, C., & Stephenson, M. C. 2021. *The credibility of corruption statistics. A critical review of ten global estimates*. Bergen: U4 Anti-Corruption Resource Centre, Chr. Michelsen Institute.
- Wei, S.-J. 1997. *Why is corruption so much more taxing than tax? Arbitrariness kills*. NBER Working Paper.
- Wu, T., Delios, A., Chen, Z., & Wang, X. 2023. Rethinking corruption in international business: An empirical review. *Journal of World Business*, 58(2): 101410.

## APPENDIX I: LIST EXPERIMENT IN MYANMAR

A list experiment is a technique that involves randomly assigning two versions of the same survey instrument to participants. This technique aids in lowering the perceived risk of detection by individual respondents, hence reducing the prevalence of social desirability error. The critical design element in a list experiment is a list of infrequent, yet non-sensitive activities, along with one other item. For the control group, this last item is a placebo, considered to have a near-zero probability of relevance in the local context. For the treatment group, this last item is a sensitive question, such as “paid informal charges to expedite application,” which measures the incidence of bribery. To shield

**Table 5** Box for Myanmar list experiment to measure size of construction bribes

Now, tell us the total cost of these activities for your business. Remember, we only need to know the total cost of all fees, not the cost of any individual fee

(Form A)	(Form B)
Paid application fees	Paid application fees
Had legal documents certified	Had legal documents certified
Paid site-inspection fees	Paid site-inspection fees
Paid informal charges to expedite application	Paid for construction project simulation by digital-modeling company
Kyat	Kyat

respondents, they mark down how many activities they have engaged in but not which activities they engaged in. When aggregating anonymized responses to the group level, analysts can calculate the frequency of bribery by comparing the difference in means between groups. Subtracting the average number of behaviors in the treatment group from the same measure in the control group results in a direct estimate of the *frequency* of bribery within the sample.

Modifications to the list experiment can help scholars measure corruption and limit exposure to the six research methods errors. We use an example of a business survey in Myanmar to illustrate these modifications (Malesky et al., 2020). This list experiment aimed to reduce social desirability error, measure different types of corruption (petty and grand), and differentiate between the *frequency* and the *scale* of corruption. Curiously, Myanmar has traditionally scored poorly for corruption, but survey work has also found that petty corruption, as measured by direct questions about informal charges, was low and declining over time (Malesky et al., 2020). Although these findings point to a possible decline in corruption, this decline might be in petty but not in grand corruption. Exact conclusions cannot be drawn because grand and collusive corruption is more opaque than petty corruption. Further, grand corruption is subject to a pronounced social desirability error due to the potential for strong punishments.

To capture the scale of corruption, respondents must not only be shielded via a list experiment, but the experiment must be altered such that the respondents are asked to report on the total cost of a group of listed activities. The control group receives a list that includes a rare but plausible activity that is not corruption, and the treatment group receives the same list except that the rare but plausible activity is replaced by corruption. Respondents are asked to report the costs of listed activities

in kyat (the local currency in Myanmar). Comparison of the treatment and control groups allows for the calculation of the average “price” of a bribe, which can also help untangle theoretical concerns about grand and petty corruption and their prevalence (Malesky et al., 2020).

The creative design element in this list experiment is the box that asks questions about the list. As not any one item from the list is ever revealed as having been done by a survey respondent, the respondents and their answers are appropriately shielded. We show a simplified version of the box from Malesky et al. (2020) in Figure A. This box is used to identify the scale of construction bribes. Respondents receive either Form A or Form B. The first three items in each form are non-sensitive, ordinary activities, randomly ordered, that are related to obtaining an operating license or a construction permit. However, Form A contains a sensitive activity related to bribery: “paid informal charge to expedite application.” Form B contains a placebo clause, such as “Paid for construction project simulation by digital-modeling company”. The placebo clause is plausible but unlikely, hence essentially taking a zero value for Form B respondents. As respondents are randomly distributed between Form A and Form B, any difference in aggregate values can be attributed to costs associated for informal charges paid to expedite applications.

See Table 5.

When analyzing the differences in aggregated responses, Form B respondents were found to have reported a mean of 404,335 kyat (USD 289) for all construction-related activities. The mean for firms answering Form A was 3,836,227 kyat (USD 2742) for all construction-related activities. These means are significantly different from one another, as the confidence intervals do not overlap. Subtracting the results of Form B from the results of Form A gives an average bribe size of 3.4 million kyat (USD 2453) per firm to obtain a construction permit.



Taking this information together, this example illustrates that the techniques required are feasible, practical, and implementable. Moreover, the list experiment technique can not only reduce social desirability error, but it can also focus on a particular form of corruption, it can address questions related to frequency and scale, and it can yield information on actual experiences on corruption. Aside from the theory development benefits of this heightened measurement acuity, list experiments provide the potential to develop better data on corruption to track its incidence and evaluate intervention attempts to reduce either its frequency or scale.

### ABOUT THE AUTHORS

**Andrew Delios** (PhD, Ivey Business School, Western University of Ontario) is Professor and Vice Dean (MSc Programs), NUS Business School, National University of Singapore. He is a Fellow of the Academy of International Business and the Asia Academy of Management. His research interests are strategy and the internationalization process, with a focus on firms operating in the Asia-Pacific region.

**Edmund J. Malesky** (PhD, Duke University) is Professor of Political Science and Director of the Duke Center for International Development. He is the Director of the Executive Council on the Southeast Asian Research Group (SEAREG) and has been the lead researcher of USAID's Vietnam Provincial Competitiveness Index project since 2005. His research focuses on the political economy of development, governance, and authoritarian institutions.

**Shu Yu** (PhD, National University of Singapore) is a research fellow at the Suzhou Industrial Park Monash Research Institute of Science and Technology and lecturer at the Monash University Suzhou Campus. Her research interests lie in corporate and international strategy, with a focus on emerging economies.

**Griffin Riddler** (Duke University) is a PhD candidate in political science. His research focuses on state-business relations, Chinese political and economic development, and historical political economy.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

*Accepted by Arjen van Witteloostuijn, Consulting Editor, 24 May 2023. This article has been with the authors for four revisions.*