Measuring Changes in Corruption over Time*

Preliminary Version: This paper is under active development. Results and conclusions may change as research progresses.

Word Count: 3010

Maya Dalton† and Justin Esarey
Wake Forest University
Department of Politics and International Affairs

April 5, 2022

Abstract

Although the most commonly used quantitative measures of corruption are highly correlated, we find that these measures do not accurately track changes in corruption within a given country over time. Many research designs rely on such within-country changes to identify causal relationships. We argue that findings based on changes in corruption within countries should be interpreted with caution. We also present a factor score measure of corruption that is less subject to measurement error.

*We gratefully acknowledge financial support for this project provided by the Wake Forest University Undergraduate Research and Creative Activities Center (URECA) and the WFU Department of Politics and International Affairs.

†Corresponding author: daltma18@wfu.edu

Can we employ existing measures of corruption to study its causes and effects using international panel data? Every measure of corruption is constructed differently, but all are intended to measure "abuse of public office for private gain" (Treisman, 2007, p. 211). All suffer from the fundamental problem that corruption is a hidden behavior whose perpetrators do not want their dealings recorded in a data set. Consequently, all of these measures are imperfect. The many potential shortcomings of these measures are well-known to corruption researchers (see Brooks et al., 2013, Chapter 2), and therefore most empirical corruption studies involve a careful consideration of the measures they employ and a detailed discussion of why some measures were highlighted. But criticisms of a measure are less concerning if a measure's construct validity is strong. As one example, aspects of the widely-used and influential Transparency International's Corruption Perceptions Index's (TI CPI, Transparency International, 2020) methodology are questionable (Lambsdorff, 2006), including whether the CPI can validly track changes in a country's corruption level over time. But we can neglect these flaws if other indicators using a different methodology report similar levels of corruption for the same country. On the other hand, if there is little correspondence between these different measures of corruption, *at most* one of the measures corresponds to the underlying concept. Determining which (if any) of the measures is correct in this situation would require such a direct understanding of the concept that an indirect measure would hardly be necessary.

Five very commonly used quantitative measures[1] of corruption—TI CPI, the World Bank's Control of Corruption governance indicator (WBGI, The World Bank Group, 2020),

[1] These five measures have exerted a strong influence on corruption research, having been cited approximately 10,800 times according to Google Scholar. We produced this figure by searching for the number of citing articles for each corruption measure; for example, for the BCI we searched *("BCI" OR "Bayesian Corruption Index") AND intitle:corruption.* Google Scholar citation counts are estimates and therefore inexact. See Appendix C for a more detailed description of each measure.

the Bayesian Corruption Index (BCI, Standaert, 2015), the Varieties of Democracy overall corruption index (V-Dem, Coppedge et al., 2021), and the International Country Risk Guide's corruption risk indicator (ICRG, The PRS Group, 2020)—are highly correlated with one another, with pairwise correlation coefficients over $\rho = 0.9$ (see Table 3 in Standaert, 2015, p. 789). However, "this is almost completely driven by their between-correlations (the correlation between the mean values for each country)" (Standaert, 2015, p. 788) and not changes in corruption within countries over time. For this reason, we may conclude that these measures are valid for cross-sectional studies of countries but not necessarily for changes in corruption within a country over time. Within-country changes are important because causal inferences are often predicated on their accurate measurement. For example, a difference-in-differences design compares changes in corruption in a treated country to changes in an untreated country over the same time period to determine the treatment's effect on corruption (Angrist and Krueger, 1999, pp. 1293-1299).

In this paper, we report that changes in corruption over time within countries are weakly correlated across different measures whether using annual corruption scores or decennial country averages. We also report that changes in corruption across different measures load similarly on a primary factor extracted via principal component analysis, but this factor explains comparatively little variance in the component scores. We conclude that these corruption measures lack construct validity for measuring change in corruption within a country over time. There are substantively meaningful differences between them not ascribable to transient measurement error. Therefore, we advise caution when interpreting findings that rely on measurements of change in corruption. Scores from a first principal component extracted from multiple panel-adjusted measures of corruption may provide a more valid measure. We provide these scores as a part of our replication data set for researchers to use and improve upon in the future.

Data

Our data set has information about 199 countries from 1980-2020 from three sources: the Quality of Government data set (QOG, Teorell et al., 2019), the V-Dem data set, and the ICRG data set.[2] The key variables in our analysis are the five influential measures of corruption we mentioned in the introduction. We re-scaled all corruption measures to range from 0 (least corrupt) to 100 (most corrupt) for ease of comparison. Some countries do not have all corruption scores available for certain years and some indicators are only available for segments of the time period.[3]

Methods

High correlation between independent measures of corruption is an indicator of their construct validity: if all measures capture the same concept, they should be closely related. Although correlation between corruption measures is strong, this is attributable to differences in corruption *between* countries and not changes in corruption within countries over time (Standaert, 2015). Hence, we need to extract differences in corruption between countries to evaluate how well indicators measure change in corruption *within* a country over time.

Consider a theoretical decomposition of corruption Y in a country i at time t over a time period $t \in \{1...T\}$:

$$Y_{it} = g_i(t) + \varepsilon_{it} \tag{1}$$

Each country i's corruption level is given by a country-specific function $g_i(t)$ that represents the trajectory of corruption over time plus an added stochastic component $\varepsilon_{it} \sim f(\mu, \sigma^2)$ with mean $\mu = 0$ and variance σ^2 that represents random influences on corruption. We speculate

[2]See Appendix A for the list of countries in our data set.

[3]See Figure 3 in Appendix B for the availability of indicators over time.

that measurements of corruption can distinguish countries' average levels of corruption from one another, but not countries' changes in corruption over time. To represent this, we rewrite equation 1 as:

$$Y_{it} = \frac{1}{T}\sum_{t=1}^{T} g_i(t) + \left(g_i(t) - \frac{1}{T}\sum_{t=1}^{T} g_i(t)\right) + \varepsilon_{it} \tag{2}$$

$$= A_i + \gamma_i(t) + \varepsilon_{it} \tag{3}$$

where A_i represents country i's average corruption over the time period $t \in \{1...T\}$ and $\gamma_i(t)$ is the de-meaned function representing its trajectory over time.

If there are any global shifts in corruption at a particular time period t, either due to measurement noise or a genuine change in the overall level of corruption worldwide, we might wish to remove these impacts from the measure as well. Removing that component is important if (as is most common) we are studying influences on corruption that vary from country to country and are not system-wide. It is also important to determine how much our corruption measures are able to distinguish country-specific changes in corruption net of overall global changes. We therefore rewrite equation 3 as:

$$Y_{it} = A_i + \frac{1}{N}\sum_{i=1}^{N} \gamma_i(t) + \left(\gamma_i(t) - \frac{1}{N}\sum_{i=1}^{N} \gamma_i(t)\right) + \varepsilon_{it} \tag{4}$$

$$= A_i + P_t + \psi_i(t) + \varepsilon_{it} \tag{5}$$

where P_t is the global average corruption rating for time t and $\psi_i(t)$ represents the remaining variance in country-specific corruption.

We can estimate the components of equation 5 with a fixed effects model:

$$y_{it} = \alpha_i + \phi_t + \omega_{it} \tag{6}$$

where y_{it} is a measure of corruption for country i in year t, α_i is the average across time in country i, ϕ_t is the average across countries in year t, and all remaining variance is in ω_{it}. Theoretically, the residual ω_{it} is equal to $\psi_i(t) + \varepsilon_{it}$, country-specific corruption plus random error. Consequently, ω_{it} is the new measure of corruption for country i at time t net of between-country variation and time-specific global shocks. α_i and ϕ_t must be defined relative to a reference category (without loss of generality); for example, ϕ_1 may be fixed at zero (i.e., swept into the mean estimate of α_i) as is typical in panel dummy variable models.

We estimate ω_{it} for the CPI, BCI, ICRG, WBGI, and V-Dem corruption measures using simple least-squares dummy variables regression.[4] We then extract the residuals from the model, ω_{it}, to create a new measure of corruption with between-country differences and worldwide time trends removed; we refer to this as a *panel-adjusted corruption measure*.

Principal components analysis (PCA) provides another window into the construct validity of a set of measures (Murphy, 2012, pp. 381-416). Suppose that we have K many observed variables; in our case, $K = 5$. In an unobserved components model, similar to the one that Kaufmann, Kraay and Mastruzzi use to construct the World Bank Governance Indicators (2011, p. 229), each observed (panel-corrected) variable ω_{itk} for country i at time t is assumed to be a composite of multiple unobserved dimensions. We are particularly interested in the dimension corresponding to corruption, ψ_{it} in equation 5. We can interpret PCA as a model of the observed scores (Bro and Smilde, 2014, pp. 2816-2817):

$$\psi_{it} = \sum_{k=1}^{K} (\delta_k \omega_{it}) + \varepsilon_{it} \tag{7}$$

That is, the unobserved level of corruption ψ_{it} is a linear combination of the observed variables. This is a close match for the theoretical definition of ω_{it} that we outlined in our discussion of panel-adjusted corruption measures. PCA chooses the vector $\vec{\delta} = \{\delta_1, \delta_2, ..., \delta_K\}$

[4]All analyses are conducted using R 4.1.0 (R Core Team, 2021), in this case with the basic `lm` function.

to maximize the variance of ψ_{it} subject to the constraint that $\sum_{k=1}^{K} \delta_k = 1$. If these measures all target the same concept, we expect this dimension to strongly determine the observed score. Specifically, if all the observed variables load onto a single dimension (in this case, ψ) that explains most of their variance, it indicates that all of the variables are measures of the same concept: corruption. If the matrix $\Omega_{NT\times K}$ is of full rank, singular value decomposition of $\Omega_{NT\times K}$ produces K many unobserved component estimates in descending order of the proportion of variance in $\Omega_{NT\times K}$ that each explains.

We therefore perform PCA analysis of our panel-adjusted measures of corruption. If all measures do not load onto a single dimension or that dimension does not explain most of the variance, the measures may :

1. target different notions of corruption;
2. be confounded by unobserved variables, or;
3. might be highly contaminated by measurement error.

We use probabilistic PCA (PPCA) to account for the missingness of some corruption measures for some country-years.[5]

Results

In this section, we report metrics of the construct validity of the CPI, BCI, ICRG, WBGI, and V-Dem corruption measures using (i) annual scores and (ii) decennial averages. For both annual scores and decennial averages, we estimate the correlation (a) among the raw corruption measures and (b) among the panel-adjusted measures created from the residuals of fixed effects models shown in equation (6). We also show PPCA results for all four cases.

Figure 1 reports correlations among raw corruption measures and panel-adjusted scores

[5]PPCA is implemented by the `pcaMethods` package (Stacklies et al., 2007).

using the residuals from a fixed effects model.[6] While the raw annual measures are highly correlated with one another (as reported by Standaert, 2015), the panel-adjusted measures are only weakly correlated (median $\hat{\rho} = 0.238$). That is, the typical corruption measure only explains about 6% of the variation in any other corruption measure once between country differences and worldwide trends in corruption are removed. The low correlation between panel-adjusted corruption measures indicates that they cannot agree on how much a country's corruption level changes over time.

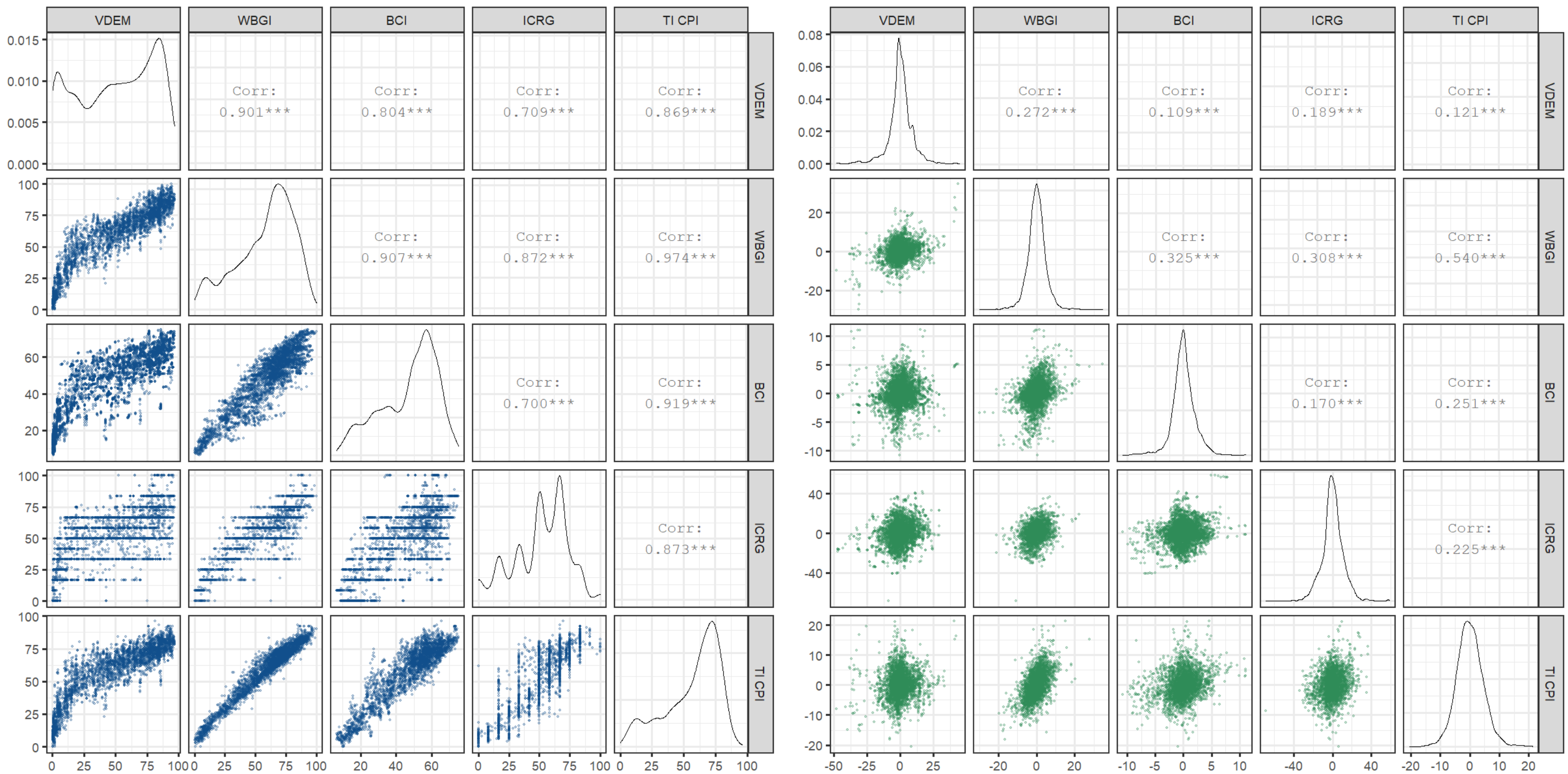
Table 1 shows factor loadings for all corruption scores on the first two principal components (PC1 and PC2) produced by PPCA; the row labeled R^2 displays the proportion of variance in corruption scores that is explained by each principal component. Among raw annual corruption scores (the first two columns in Table 1), about 90% of the variance in corruption measures is accounted for by a single dimension (PC1); all corruption scores load positively on this dimension.[7] Although the dimensions extracted by PPCA do not have an intrinsic interpretation, the fact that all corruption measures load positively on a single dimension suggests that they all map onto a single concept: *corruption*.

For the panel-adjusted measures of corruption, PPCA identifies a PC1 dimension (shown in the second two columns of Table 1) with factor loadings extremely similar to the PC1 dimension for raw corruption scores. However, PC1 only explains 39.5% of variance in the within-country change measures of corruption as compared to 90% for the raw measures. The second dimension extracted by PPCA (PC2) explains much more variance for the panel-adjusted measures compared to the second PPCA dimension for raw corruption scores. This finding suggests that corruption measures *do* track a common component of

[6] For more detailed correlation figures between raw measures and fixed effect residual scores, see Table 2 in Appendix D.

[7] We multiplied some factor loading matrices by -1 to place all first principal component loadings in the same direction.

Figure 1: **The correlation among raw and panel-adjusted measures of corruption**



(a) Raw Measures

(b) Panel-Adjusted Measures

Figure 1a depicts the correlation between pairs of raw measures of corruption (named at the top and right panels). Figure 1b depicts the correlation between panel-adjusted measures. The stars indicate statistical significance: $(* = p < 0.1, ** = p < 0.05, *** = p < 0.01)$.

change in corruption within a country over time but with much less accuracy than they track between-country differences in corruption. Furthermore, there are other (substantively unidentified) common factors shared by these measures that explain a considerable portion of their variance. For example, PC2 explains over 21% of the variance in corruption for the panel-adjusted scores.

Table 1: **Factor Loadings and R^2 for Principal Components**

	annual raw		annual panel-adjusted	
	PC1	PC2	PC1	PC2
VDEM	0.473	-0.362	0.580	-0.767
WBGI	0.463	-0.023	0.436	0.174
BCI	0.451	-0.369	0.473	0.570
ICRG	0.403	0.856	0.372	0.065
TI CPI	0.443	0.006	0.334	0.227
R^2	0.900	0.050	0.395	0.218

The factor loadings on the first two principal components for all corruption scores. The R^2 row displays the proportion of variance in corruption scores that is explained by each principal component.

Conclusion

Based on our findings, we advise researchers to use caution when change in corruption over time within countries is a key independent or dependent variable. The suite of commonly used corruption indicators do not agree when measuring changes in corruption within a given country over time, even when comparing long-run average changes.

The substantive importance of this disagreement is easily illustrated via example. Consider Figure 2, which shows panel-adjusted corruption measurements over time for China (panel 2a) and South Africa (panel 2a). We have standardized each measure (by subtracting the mean and dividing by the standard error) within each country. While the measures clearly share commonalities, from year to year they often disagree on the magnitude and direction of change. For example, in China the ICRG indicates a relatively stable corruption level between 2005 and 2018. But the other measures (and our PCA score) indicate a sharp decline in corruption. In South Africa between 2010 and 2020, two measures stay relatively stable, one indicates a dramatic increase in corruption, and the other indicating an initial increase followed by a sharp *decrease* in corruption.

This example and our analysis shows that changes in corruption measures *do* share a common factor. But this factor explains relatively little of the variance in each component measure. Concordantly, we believe that the best option at present is to be cautious when interpreting results that depend on within-country changes in corruption. A result should be replicated using many measures of corruption to assess its robustness. If results diverge, this should be reported and the theoretical importance of this divergence should be assessed. As an alternative, it is possible for researchers to (a) extract a common factor from multiple corruption measures using probabilistic principal components analysis of panel-adjusted corruption measures and (b) use the resulting scores as a measure of (change in) corruption within countries. We have provided these scores for the international system between 1980 and 2020 as a part of the replication material for this paper to make them easy for future researchers to study, criticize, and use.

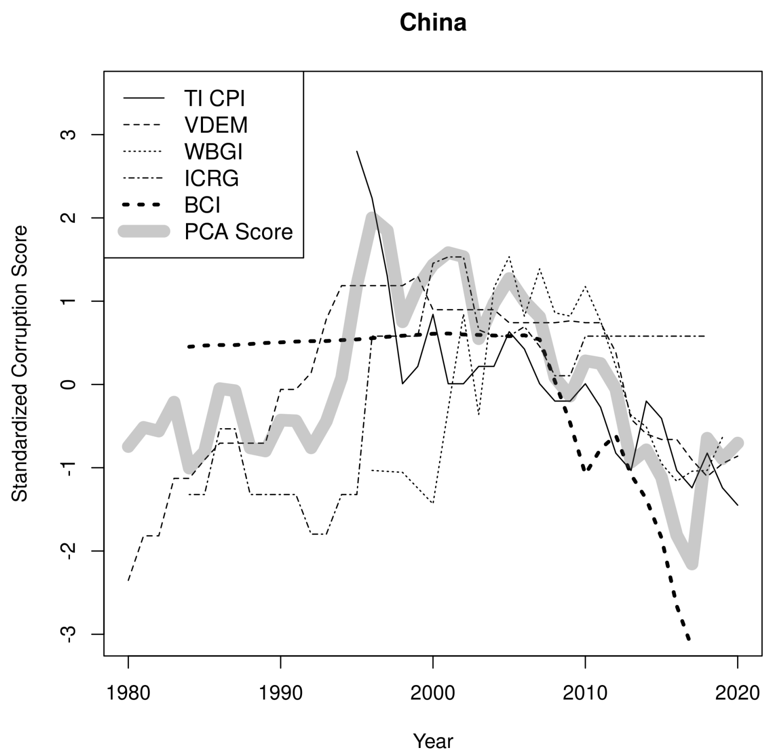
References

Angrist, Joshua D. and Alan B. Krueger. 1999. Empirical strategies in labor economics. In *Handbook of Labor Economics*. Vol. 3 Elsevier pp. 1277–1366.

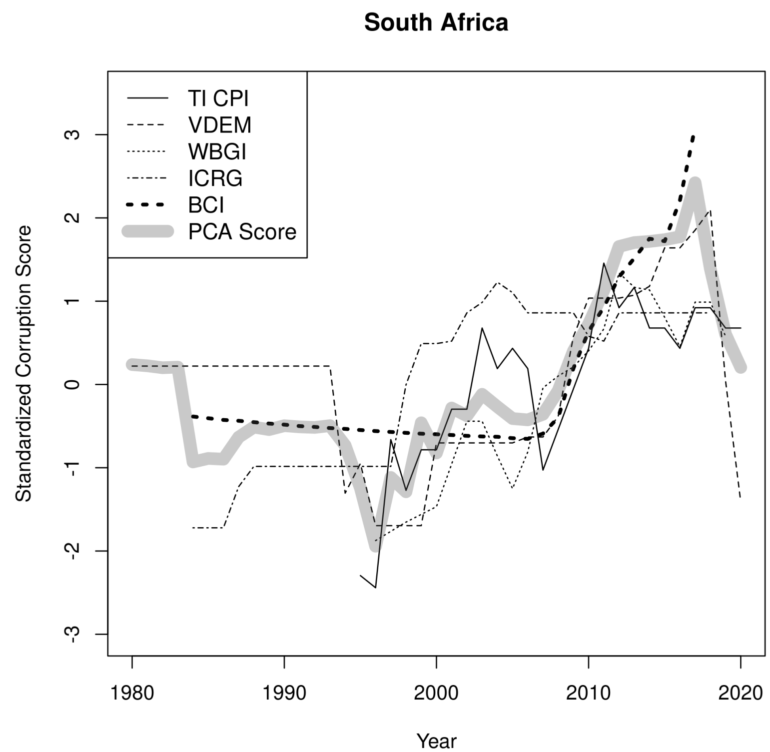
Bro, Rasmus and Age K. Smilde. 2014. “Principal component analysis.” *Analytical Methods* 6:2812–2831.

Brooks, Graham, David Walsh, Chris Lewis and Hakkyong Kim. 2013. *Preventing Corruption: Investigation, Enforcement, and Governance*. Palgrave Macmillan UK.

Figure 2: **Two examples of divergent corruption measures**



(a) China



(b) South Africa

Figure 2a depicts the trajectory of six standardized corruption measures (the five observed panel-adjusted variables and our extracted PCA score) from 1980 to 2020 for China. Figure 2b depicts the trajectory of the same standardized corruption measures from 1980 to 2020 for South Africa.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen et al. 2021. *V-Dem Codebook v11.1*. Varieties of Democracy (V-Dem) Project. URL: https://www.v-dem.net/en/data/data/v-dem-dataset-v111/.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2011. "The Worldwide Governance Indicators: Methodology and Analytical Issues." *Hague Journal on the Rule of Law* 3(2):220–246.

Lambsdorff, Johann Graf. 2006. Measuring Corruption-The Validity and Precision of Subjective Indicators (CPI). In *Measuring Corruption*, ed. Charles J. G. Sampford, Arthur Shacklock, Carmel Connors and Fredrik Galtung. Burlington, VT: Ashgate Publishing, Ltd. pp. 81–99.

Murphy, Kevin P. 2012. *Machine Learning: A Probabilistic Perspective*. MIT Press. URL: https://www.cs.ubc.ca/~murphyk/MLbook/.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: https://www.R-project.org/.

Stacklies, Wolfram, Henning Redestig, Matthias Scholz, Dirk Walther and Joachim Selbig. 2007. "pcaMethods – a Bioconductor package providing PCA methods for incomplete data." *Bioinformatics* 23:1164–1167.

Standaert, Samuel. 2015. "Divining the level of corruption: A Bayesian state-space approach." *Journal of Comparative Economics* 43(3):782–803. URL: https://doi.org/10.1016/j.jce.2014.05.007.

Teorell, Jan, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon and Richard Svensson. 2019. "The Quality of Government Dataset, version Jan19." URL: http://www.qog.pol.gu.se.

The PRS Group. 2020. "International Country Risk Guide: Corruption Risk Measure." URL: https://www.prsgroup.com/explore-our-products/international-country-risk-guide/.

The World Bank Group. 2020. "Worldwide Governance Indicators." URL: https://datacatalog.worldbank.org/dataset/worldwide-governance-indicators.

Transparency International. 2020. "Corruption Perceptions Index Overview." URL: https://www.transparency.org/en/cpi/2020/index/nzl.

Treisman, Daniel. 2007. "What Have We Learned About the Causes of Corruption from Ten Years of Cross-National Empirical Research?" *Annual Review of Political Science* 10(1):211–244. URL: http://www.annualreviews.org/doi/abs/10.1146/annurev.polisci.10.081205.095418.

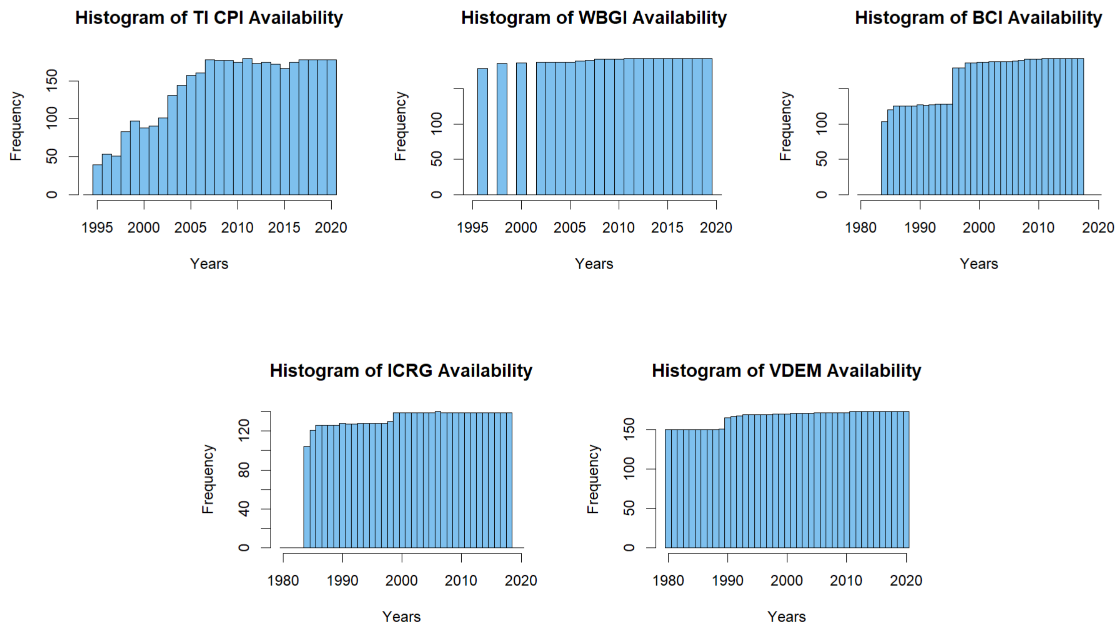
Appendices

A List of Countries with Available Data from 1980-2020

Afghanistan, Albania, Algeria, Andorra, Angola, Antigua and Barbuda, Azerbaijan, Argentina, Australia, Austria, Bahamas, Bahrain, Bangladesh, Armenia, Barbados, Belgium, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Belize, Solomon Islands, Brunei, Bulgaria, Myanmar, Burundi, Belarus, Cambodia, Cameroon, Canada, Cape Verde, Central African Republic, Sri Lanka, Chad, Chile, China, Taiwan, Colombia, Comoros, Congo, Congo, Democratic Republic, Costa Rica, Croatia, Cuba, Cyprus (1975-), Czechoslovakia, Czech Republic, Benin, Denmark, Dominica, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Ethiopia (-1992), Ethiopia (1993-), Eritrea, Estonia, Fiji, Finland, France (1963-), Djibouti, Gabon, Georgia, Gambia, Germany, Germany, East, Germany, West, Ghana, Kiribati, Greece, Grenada, Guatemala, Guinea, Guyana, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Cote d'Ivoire, Jamaica, Japan, Kazakhstan, Jordan, Kenya, Korea, North, Korea, South, Kuwait, Kyrgyzstan, Laos, Lebanon, Lesotho, Latvia, Liberia, Libya, Liechtenstein, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia (1966-), Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Moldova, Montenegro, Morocco, Mozambique, Oman, Namibia, Nauru, Nepal, Netherlands, Vanuatu, New Zealand, Nicaragua, Niger, Nigeria, Norway, Micronesia, Marshall Islands, Palau, Pakistan (1971-), Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Guinea-Bissau, Timor-Leste, Qatar, Romania, Russia, Rwanda, St Kitts and Nevis, St Lucia, St Vincent and the Grenadines, Sao Tome and Principe, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Singapore, Slovakia, Vietnam, Slovenia, Somalia, South Africa, Zimbabwe, Spain, South Sudan, Sudan (2012-), Sudan (-2011), Suriname, Eswatini (former Swaziland), Sweden, Switzerland, Syria, Tajikistan, Thailand, Togo, Tonga, Trinidad and Tobago, United Arab Emirates, Tunisia, Turkey, Turkmenistan, Tuvalu, Uganda, Ukraine, North Macedonia, USSR, Egypt, United Kingdom, Tanzania, United States, Burkina Faso, Uruguay, Uzbekistan, Venezuela, Samoa, Yemen, Serbia and Montenegro, Zambia

B Corruption Measure Availability by Year

Figure 3: **Availability of Corruption Measures by Year**



The availability of each of the five corruption measures during different time periods, ranging from 1980 to 2020.

C Descriptions of Frequently-Used Corruption Measures

The descriptions in this section are identical to descriptions in the online appendix of Esarey and Dalton (2021); these papers were written at the same time using (some of) the same variables.

1. *Transparency International's Corruption Perceptions Index (CPI)*[8]
   The CPI is an extremely influential indicator of corruption widely used by scholars and policymakers.[9] It is constructed from averaging at least three (but as many as thirteen) different corruption scores taken from perception-based surveys and assessments of corruption in a given country. The CPI targets corruption in the public sector within a country and compiles relevant data from multiple, independent sources. The CPI standardizes the corruption scores from these sources to the same scale, then averages the scores. Finally, the standard error and confidence interval for each country's CPI value is calculated to account for any variation in the sources. The CPI ranges from 0 (most corrupt) to 100 (least corrupt), and is available from 1995-2020.
2. *World Bank Group's Worldwide Governance Indicators (WBGI)*[10]
   The WBGI is created from 30 data sources from a variety of surveys, organizations, and governments. It utilizes a Unobserved Components Model (UCM) to construct six aggregated indicators of governance and estimate margins of error for each indicator. Of the six indicators, our interest is in their measure of *control of corruption*, defined by Kaufmann, Kraay and Mastruzzi (2010, p.4) as "the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as 'capture' of the state by elites and private interests". The WBGI ranges from -3 (least control over corruption - highly corrupt) to 3 (most control over corruption - least corrupt), and is available for 1996, 1998, 2000, and 2002-2020.
3. *Bayesian Corruption Index (BCI)*[11]
   The BCI is an index of perceived overall corruption (abuse of public power for private gain) within a country. It is constructed from 17 different surveys from countries' inhabitants, business executives, and governments. The BCI expands upon the number of sources used by the WBGI and CPI and is available over a larger time span than either of these two measures, but the measurement models used by the BCI and WBGI are

[8] Information about the CPI has been paraphrased from Transparency International (2016) and Transparency International (2020)

[9] According to Galtung (2006, p. 106), "The impact of the CPI has been considerable. It has been credited as a factor that gave the issue of corruption 'greater international prominence' (Florini, 1998).... The CPI has facilitated a qualitative shift in the journalistic writing and public discourse on corruption.... This interest and awareness of the CPI extends well beyond the business and financial press."

[10] Information about the WBGI has been paraphrased from Kaufmann, Kraay and Mastruzzi (2010).

[11] Information about the BCI has been paraphrased from Standaert (2015).

broadly similar. Unlike the WBGI, the BCI's measurement model accounts for variation over time to avoid discrepancies in corruption measurements and prevent selection bias. The BCI ranges between 0 (least corrupt) to 100 (most corrupt) in countries and is available from 1984 onward.

4. *Political Risk Service's International Country Risk Guide (ICRG)*[12]
   The ICRG provides political, economic, and financial risk ratings to inform businesses about potential risks to their firms when operating within certain countries. It aims to assess how much corruption within the political system can threaten foreign investment and political stability. It measures corruption through the prevalence of bribery, patronage, nepotism, extortion, suspicious ties between business and politics, and related phenomena. The measurement ranges between 0 (low political risk from corruption) and 6 (high political risk from corruption), is a source included in the CPI, and available for 1994 onward.

5. *Varieties of Democracies (V-Dem)*[13]
   The V-Dem project as a whole constructs 470 democracy measures created from subjective, expert-led assessments that score how well governments are performing relating to democratic ideals. One of their products is a measure of overall corruption in a country-year. This composite measure is created from averaging four other sub-indicators of corruption: (i) the public sector corruption index, (ii) the executive corruption index, (iii) a measure of legislative corruption, and (iv) a measure of judicial corruption. These four measures are in turn created from expert assessments of corruption in the corresponding sector of government. The resulting composite measure of overall corruption ranges from 0 to 1, with 0 indicating low corruption, and is available from 1980 to 2020.

[12] Information about the ICRG has been paraphrased from The PRS Group (2020)

[13] Information about the V-Dem has been paraphrased from Coppedge et al. (2021)

D Annual Correlation Table

Table 2: **Correlation Among Corruption Measures**

Measure	Data Type	VDEM	WBGI	BCI	ICRG	TI CPI
VDEM	raw score, annual	1.000	0.901	0.804	0.709	0.869
	panel-adjusted, annual	1.000	0.272	0.109	0.189	0.121
	panel-adjusted, decade avg.	1.000	0.381	0.144	0.272	0.202
WBGI	raw score, annual	0.901	1.000	0.907	0.872	0.974
	panel-adjusted, annual	0.272	1.000	0.325	0.308	0.540
	panel-adjusted, decade avg.	0.381	1.000	0.392	0.417	0.573
BCI	raw score, annual	0.804	0.907	1.000	0.700	0.919
	panel-adjusted, annual	0.109	0.325	1.000	0.170	0.251
	panel-adjusted, decade avg.	0.144	0.392	1.000	0.228	0.246
ICRG	raw score, annual	0.709	0.872	0.700	1.000	0.873
	panel-adjusted, annual	0.189	0.308	0.170	1.000	0.225
	panel-adjusted, decade avg.	0.272	0.417	0.228	1.000	0.259
TI CPI	raw score, annual	0.869	0.974	0.919	0.873	1.000
	panel-adjusted, annual	0.121	0.540	0.251	0.225	1.000
	panel-adjusted, decade avg.	0.202	0.573	0.246	0.259	1.000
	# of country-years	6802	3977	5639	4662	3650

Correlation among corruption measures for raw annual scores, panel-adjusted annual scores, and panel-adjusted decennial average scores. VDEM = the Varieties of Democracy political corruption index; WBGI = the World Bank Governance Indicators Control of Corruption measure; BCI = the Bayesian Corruption Index; ICRG = the International Country Risk Group's corruption risk measure; TI CPI = Transparency International's Corruption Perception index.

References

Coppedge, Michael, John Gerring, Carl Henrik Knutsen et al. 2021. *V-Dem Codebook v11.1*. Varieties of Democracy (V-Dem) Project. URL: https://www.v-dem.net/en/data/data/v-dem-dataset-v111/.

Esarey, Justin and Maya Dalton. 2021. "The Changing Relationship between Gender and Corruption." Working Paper. URL: http://www.justinesarey.com/Corruption_and_Gender_Over_Time__Dalton_and_Esarey.pdf.

Florini, Ann. 1998. "The End of Secrecy." *Foreign Policy* (111):50–63.

Galtung, Fredrik. 2006. Measuring the Unmeasurable: Boundaries and Functions of (Macro) Corruption Indices. In *Measuring Corruption*, ed. Charles J. G. Sampford, Arthur Shacklock, Carmel Connors and Fredrik Galtung. Burlington, VT: Ashgate Publishing, Ltd. pp. 101–130.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues*. SSRN. URL: http://ssrn.com/paper=1682130.

Standaert, Samuel. 2015. "Divining the level of corruption: A Bayesian state-space approach." *Journal of Comparative Economics* 43(3):782–803. URL: https://doi.org/10.1016/j.jce.2014.05.007.

The PRS Group. 2020. "International Country Risk Guide: Corruption Risk Measure." URL: https://www.prsgroup.com/explore-our-products/international-country-risk-guide/.

Transparency International. 2016. "Technical Methodology Note." URL: https://bit.ly/2Iu9IrU accessed 6/20/2019.

Transparency International. 2020. "Corruption Perceptions Index Overview." URL: https://www.transparency.org/en/cpi/2020/index/nzl.