Measuring Transparency*

James R. Hollyer†1, B. Peter Rosendorff[2], and James Raymond Vreeland[3]

[1]Department of Political Science, University of Minnesota

[2]Wilf Family Department of Politics, New York University

[3]Mortara Center for International Studies, Georgetown University

December 2013

Abstract

Transparency is often viewed as crucial to government accountability, but its measurement remains elusive. This concept encompasses many dimensions, which have distinct effects. In this paper, we focus on a specific dimension of transparency: governments' collection and dissemination of aggregate data. We construct a measure of this aspect of transparency, using an item response model that treats transparency as a latent predictor of the reporting of data to the World Bank's World Development Indicators. The resultant index covers 125 countries from 1980-2010. Unlike some alternatives (e.g., Freedom House), our measure – the HRV Index – is based on objective criteria rather than subjective expert judgments. Unlike newspaper circulation numbers, HRV reflects the dissemination of credible content – in that it has survived the World Bank's quality control assessment. In a validation exercise, we find that our measure outperforms newspaper circulation as a predictor of Law and Order and Bureaucratic Quality as measured by the ICRG, particularly in autocracies. It performs as well as newspaper circulation in predicting Corruption. These findings suggest that data dissemination is a distinct, and politically relevant, form of transparency.

*We would like to thank Neal Beck, Adam Bonica, Lawrence Broz, Christina Davis, Jennifer Gandhi, Michael Gilligan, Andrew Little, Jonathan Nagler, Megumi Naoi, William Skorupski, Randall Stone, Johannes Urpelainen and participants in the 2011 MPSA Panel on New Methodologies and Data for IPE, the NYU Workshop on Bayesian Methods, the 2012 ISA Panel on Transparency Information and Firms in a Global Economy, and the Stanford University Methods Workshop for helpful comments and suggestions. We would also like to thank the Minnesota Supercomputing Institute for access to their resources. All remaining errors are our own. Replication materials for this paper are available on the Dataverse site for this paper, http://dx.doi.org/10.7910/DVN/24274. We will also upload all replication files, our index, and related work to http://HRVtransparency.org.

†Corresponding Author: 1414 Social Sciences, 267 19th Ave. S, Minneapolis, MN 55455. jhollyer@umn.edu

Transparency, as broadly defined, relates to the full flow of information within a polity. A lengthy literature stresses the importance of the provision of information to the populace to ensure the accountability of government, and thus emphasizes the role of transparency as a determinant of government performance (see, for instance Adserà, Boix and Payne, 2003; Besley and Burgess, 2002).

This broad notion of transparency may encompass many different forms of information transmission. Common proxies for transparency measure aspects of the media market, the presence/absence of freedom of information laws (FOILs), or even political institutions such as democracy. A smaller portion of the literature considers an alternative dimension of transparency: the government's collection and dissemination of aggregate economic data.

In this paper, we introduce a novel measure of this last dimension of transparency. To construct this measure, we rely on a Bayesian item response (IRT) model which treats transparency (along this dimension) as a latent predictor of the reporting or non-reporting of data in the World Bank's World Development Indicators (WDI) data series. Our model analyzes 240 items corresponding to 240 variables consistently collected by the WDI over time. Since the WDI obtains its data from other international agencies that, in turn, obtain their data from national statistical offices, our measure is a valid indicator of governments' efforts to collect and disseminate economically relevant information. Moreover, because the WDI omits data considered 'questionable', our index reflects the collection and dissemination of credible information.[1]

Our index thus focuses on a specific dimension of transparency that has received only episodic treatment in the literature on information transmission.[2] This measure is based on objective criteria, rather than on subjective expert judgments. Moreover, our reliance on the presence or absence of data from a long-running data series provides greater coverage than many commonly used alternatives. We construct transparency measures for 125 countries covering a period from 1980-2010 – a total of 3,875 observations. The resultant index is measured at the interval level – by construction, a unit change in our scale has a consistent meaning across all index values.

Elsewhere (Hollyer, Rosendorff and Vreeland, 2013*b*), we consider the relationship between – and offer

[1]See the World Bank statements about World Development Indicators: `http://data.worldbank.org/about/data-programs`, accessed March 7, 2011. In some instances, these data are omitted by the WDI. In others, the international agencies that provide information to the WDI weed out questionable data earlier in the collection process.

[2]Notable exceptions inlcude Islam (2006) who focuses on the speed with which data are reported; Bueno de Mesquita et al. (2003), Rosendorff and Vreeland (2006), Lebovic (2006), and Williams (2009), who examine aspects of data dissemination; and Stone (2008), who treats data dissemination as a measure of state capacity.

a taxonomy of – the various facets of transparency. We acknowledge that our data dissemination index does not address all facets of transparency, such as media freedom or the openness of the policy-making process. Data dissemination, we argue, is likely to play a particularly critical role in ensuring government accountability with respect to policies when (1) these policies affect a wide portion of the population, (2) when the public is uninformed of the consequences of specific policy choices, and (3) when factors other than government decision-making affect policy outcomes – e.g., with respect to economic policy-making. Bueno de Mesquita et al. (2003) and Hollyer, Rosendorff and Vreeland (2013*b*) further contend that this form of transparency may facilitate collective action – such as protest or revolt – to discipline governments when access to the ballot box is unavailable or constrained.

We thus offer a precise measure of a limited but important component of transparency that has important implications for government accountability. We emphasize that our measure does not capture all the dimensions of transparency, that accountability itself is complex and influenced by many factors. Moreover, data dissemination is related to other factors beyond government accountability, notably a country's level of economic development. Still, our measure is crisply defined, reproducible, meaningful, and important as both an intrinsically interesting outcome variable and as a potential explanatory variable. We encourage scholars to use it – along with other appropriate measures of other facets of transparency – to test various hypotheses about the relationships between transparency and accountability, governance, and democracy. These data are available for public use at ...

In what follows, we first discuss existing measures of transparency. We contend that the bulk of the literature ignores an important component of transparency: governments' collection and dissemination of credible aggregate data on policy (particularly economic policy) outcomes. Those works that have considered this form of transparency are subject to methodological shortcomings which we seek to address with this new index. Section 2 considers whether data dissemination reflects state capacity or the willingness to disclose data, and concludes that both are necessary – and neither sufficient – to ensure high levels of disclosure. We present our empirical model in Section and demonstrate the measure's value added relative to simpler approaches in Section 4. Section 5 considers the face validity of the resultant measures, and Section 6 conducts a more extensive validation exercise in which we replicate the results of Adserà, Boix and Payne (2003), using both our measure and their original index of transparency. Section 7 concludes.

1 Existing Measures of Transparency

A large literature in political science relates transparency – broadly conceived of as the flow of information – to government accountability. These theories predominantly build upon retrospective voting models (Barro, 1973; Ferejohn, 1986), in which citizens discipline the government by the threat of removal from office. Information is modeled as improving the precision (reducing the variance) with which the public observes either the outcome of government decisions or the policy choices adopted by the government (Besley, 2006). These studies typically conclude that improvements in the flow of information to the citizenry enable citizens to adopt superior strategies for disciplining their leaders, thereby improving government performance. These theoretical findings have been given empirical support in both cross-national and sub-national analyses.

These studies vary substantially in their operationalization of transparency. In cross-national studies, the most commonly used proxies capture two distinct aspects of the media market – notably the freedom of the press, and the penetration of the press. Freedom House's Freedom of the Press index is the most frequently used measure of the former concept. This index uses subjective expert rankings to place each country year on a 0-100 scale measuring (1) the laws and regulations that constrain media content, (2) the degree of political control over the media, and (3) the structure of media ownership.[3] Numerical rankings are available from 1994 and coverage extends to 197 countries in 2013.[4] For instance, Brunetti and Weder (2003) examine the relationship between corruption and the freedom of the press and find that perceived corruption declines as press freedom rises in a cross-section of countries.

Other major indexes focus on media penetration. The most commonly used of these measures focuses on daily newspaper circulation per 1,000 residents and is made available by the World Bank. These data are available in five year increments prior to 1995 and annually thereafter. Adserà, Boix and Payne (2003) use this index to demonstrate the transparent democracies receive significantly higher governance scores than non-transparent democracies. Others, such as Besley and Burgess (2002), use analogous measures at the subnational level (see also, Reinikka and Svensson, 2003).

Both these measures come with substantial drawbacks. The Freedom House scores are coded sub-

[3] See `http://freedomhouse.org/templace.cfm?page=533`.

[4] Observations are available from 1980 based on a trichotomous $\{free,\ not\ free,\ paritaly\ free\}$ measure.

jectively and are arbitrarily constrained to a 0-100 scale. Both measures are have limited temporal and/or cross-sectional coverage. Moreover, as a measure of transparency each has substantial flaws: The press may be free from government intervention, but only reach a limited audience. Somalia, for example, receives a Freedom of the Press score roughly two standard deviations above the mean value in 1994, even though the press' reach was limited. Conversely, the press may be pervasive because it is dominated and distributed by a powerful state. Kuwait has newspaper circulation figures more than two standard deviations above the mean in 1995, but the press faced constraints on what it could report. Finally, and most significantly from our perspective, neither a free nor a pervasive press is sufficient to ensure the dissemination of credible aggregate economic information.

Other measures of transparency focus on political institutions, or legal provisions that give access to internal government information used in policy-making or assessments of policy-implementation. Broz (2002) argues that democracies have an inherently more open decision-making process than autocracies, and thus are less reliant on currency pegs as a signal to investors. Turning to more specific legal provisions, Islam (2006) finds that countries with FOILs have better governance than those without (on the passage of FOILs see, Berliner, 2011). Many studies, examine the effects of the release of audits of governmental agencies on governance outcomes (e.g. Björkman and Svensson, 2009; Chong et al., 2010; Di Tella and Schargrodsky, 2003; Ferraz and Finan, 2008).

None of these measures, however, captures the role of the disclosure of aggregate economic information to the public. Such information is unlikely to be provided by the press since (1) its collection entails large fixed costs, and (2) information is subject to a public goods problem (Rodrik, 1995). Moreover, such information is likely to prove critical in assessing government performance in a wide variety of settings. In particular, such information will help enable citizens whether fluctuations in their personal economic experiences are the result of government behavior or other economic shocks (for a related argument, see Duch and Stevenson, 2008).

Scholars have paid relatively scant attention to the collection and reporting of aggregate economic data. Two notable exceptions are Islam (2006) and Williams (2009). Islam (2006) constructs an index based on the speed with which governments report data to the WDI, to the International Monetary Fund's

International Financial Statistics (IFS), and other sources.[5] She examines 11 indicators drawn from across these series and assigns scores based on the degree of deviation from the 'desirable' frequency with which such data should be reported. These scores are then added together to construct an aggregate measure of data reporting frequency.

Williams (2009), like this paper, focuses specifically on the reporting of data. His measure is constructed based on the reporting of data to the WDI and IFS, and relies on the fraction of these variables reported by a given country (controlling for the influence of time). (For a similar approach see, Hollyer, Rosendorff and Vreeland 2011.)

In the work that follows, we construct an index that, like Islam's, focuses on data dissemination. Unlike Islam, we focus on the simple presence or absence of data from the WDI, rather than on the speed with which data are reported. Unlike Williams (2009) and Hollyer, Rosendorff and Vreeland (2011), we do not rely simply on the fraction of variables reported to the WDI – which implicitly assumes that the reporting of one variable is equivalent to the reporting of any other. We construct our index through an IRT model that ensures minimal information loss from collapsing a 240 dimension observation into a single dimension representation. Our index (henceforth the HRV index), relies on the reporting of a broad range of variables from across the WDI and offers substantial cross-sectional and longitudinal coverage (125 countries over a 31 year period).

We contend that the HRV Index is a valid measure of the degree of data dissemination by the government. Like other measures of transparency, ours is limited to a specific facet of information transmission. It thus has strengths and weaknesses, and its appropriateness depends on the theory being tested. Our index, we argue, has a key property: that of aggregation. Aggregate data indicate the effectiveness of government policy better than the experiences of individual citizens. When factors other than government actions influence the welfare of citizens, they cannot deduce the appropriateness of policies based on their individual outcomes alone. As government policies grow less predictive of any given citizen's welfare, the need for aggregate data to ensure government accountability rises.

[5] Islam (2006) also examines the role of an alternative form of transparency – the operation of FOILs.

2 Missing Data as an Element of Transparency

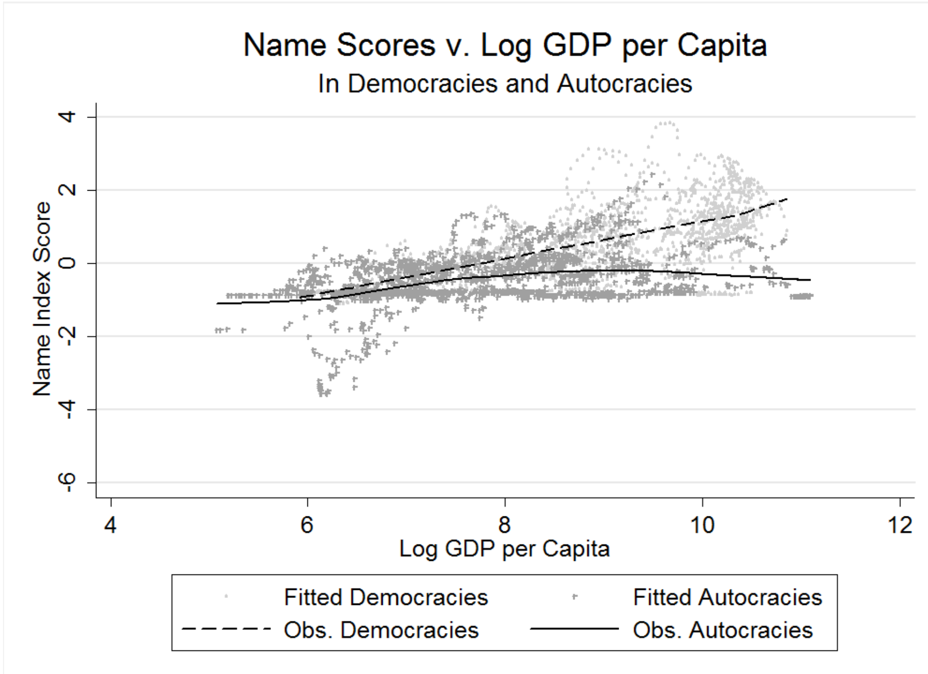
Our concept of theoretical interest is therefore disclosure of aggregate economic information to the public. Our empirical measure of this concept is a function of the missingness/non-missingness of data from the WDI.

Existing studies disagree regarding the interpretation of measures of data disclosure. Bueno de Mesquita et al. (2003), Hollyer, Rosendorff and Vreeland (2011) and Williams (2009) focus on missing data as a reflection of governments' willingness to disseminate data. Contrastingly, Stone (2008) uses missing data as a proxy for state capacity, assuming that missing data result from the inability of a government to collect and publish information. We suggest that these two explanations are non-exclusive: both the government's willingness and its capacity to collect and distribute data are necessary, but neither is sufficient, to ensure transparency.

Evidence for this claim can be seen in Figure 1. This figure plots normalized HRV index scores (on the y-axis) against the natural logarithm of *per capita* GDP (measured in constant 2005 US dollars from the Penn World Table 7.1 Heston, Summers and Aten 2012) for both democracies and autocracies. *Per capita* GDP is an often-used proxy for state capacity (Fearon and Laitin, 2003). Locally weighted scatterplot smoothing (lowess) curves are plotted for observations of both regime types. Several aspects of this relationship are immediately apparent, which serve to suggest that data dissemination is not simply a reflection of state capacity: (1) The lowess curve for democracies is consistently above that of autocracies, indicating that democratic regimes are more transparent, controlling for income. (2) The slope of the lowess curve is steeper for democracies than for autocracies. Although low income democracies report at similar rates as low income autocracies – both are constrained by a lack of capacity – democracies grow significantly more likely to report as income rises. This relationship is likely attributable to a greater *incentive* on the part of democracies to report. (3) Contrary to what one would expect if dissemination were simply a function of capacity, the relationship between transparency levels and income in autocracies is largely flat and shows signs of a slight non-monotonicity. Reporting rises in income for relatively poor autocracies, but this relationship breaks down – and there is some evidence that disclosure even declines in income – for more wealthy autocracies.

We present further evidence for the importance of political will in the decision to disclose data in the

Figure 1: HRV Scores versus Log GDP



Scatter plots, and fitted lowess lines, of standardized HRV scores against the natural log of GDP *per capita*. All HRV index scores have been standardized by subtracting the index mean and dividing by the index standard deviation. Autocracies are plotted as dark gray points, democracies as light gray points. The dashed line is the lowess curve fit to democratic observations. The solid line is the lowess curve fitted to autocratic observations.

ensuing section. The most discriminating variables in our model primarily relate to politically sensitive aspects of the economy – notably trade and investment figures. The least discriminating primarily relate to politically uncontroversial concerns – population size, for example. These results suggest that disclosure is a political decision, not simply a reflection of bureaucratic capacity.

Note that, from the point of view of many citizens, it may not matter whether a paucity of data results from state willingness or capacity. In the absence of data disclosure, this citizen will be less able to draw accurate inferences about government behavior. Regardless of whether her government was *unable* or *unwilling* to collect and disclose data, she is equally uninformed.

Nonetheless, one should take care in using our index. Because it is likely to be correlated with state capacity – and because variation in the levels of capacity may help determine government performance – it is possible that the technical abilities and competence of a government's bureaucracy may act as a confound in the relationship between governance outcomes and transparency. Other potential confounding variables may include a country's political institutions, level of exposure to international trade, and degree of participation in international organiztions. All of these factors may jointly determine transparency and government performance. Researchers employing our measure would therefore be well-advised to take measures to control for these confounding influences in empirical specifications.

3 The Empirical Model

We measure governments' collection and disclosure of data directly, by relying on the presence or absence of reported values from the WDI. The World Bank assembles these data from information provided by other international organizations – for instance, the International Monetary Fund (IMF) and International Labor Organization (ILO). These organizations, in turn, obtain information from national statistical offices. In some instances, the World Bank codes observations in the WDI as missing because the information provided by national agencies is deemed to be questionable. Our measure thus reflects the disclosure of credible information by national statistical agencies. As such, it possesses a high degree of *content validity* (Carmines and Zeller, 1979) with regards to the concept of theoretical interest – the collection and dissemination of aggregate economic data.[6]

To be more precise, we treat *transparency* as a latent term predicting the presence or missingness of data on 240 measures drawn from the WDI. We obtain estimates of this transparency term through the use of a Bayesian IRT model.[7]

Our index is thus a summary measure of the tendency to report these 240 variables to the WDI. For this index to be a valid indicator of a government's tendency to release aggregate economic data, these variables must be representative of the type of data collected by governments more generally and must not be subject to any arbitrary reporting procedures. In short, the quality of our index is a direct function of the data that serves as an input to our model. We therefore discuss the procedure for selecting these

[6] We should note, however, that while data quality is a continuous concept, the WDI's decision to omit or report data is discrete. Consequently, missingness is not fully indicative of data quality.

[7] All materials necessary to replicate our measurement index, figures and tables from this paper, and robustness checks from the appendix are available on the Dataverse site for this paper, http://dx.doi.org/10.7910/DVN/24274. We will also upload all replication files, our index, and related work to http://HRVtransparency.org.

variables in some detail below.[8]

It is unlikely that all 240 items are necessary for the construction of this index. In the appendix, we present estimates of the discrimination scores for each variable from the WDI used in the index. The discrimination parameters on many of these items – such as the crude birth rate, the percentage of the population that is female, the percentage of the population aged 0-14 – are statistically indistinguishable from zero. This implies that they contribute little to index values and could likely be safely excluded from the model. One of the virtues of the IRT approach, however, is that the importance of each item is a quantity to be estimated, rather than assumed. We therefore prefer to include too many rather than too few variables in our model, and let the data speak as to their importance.[9]

The 240 variables used in our model are drawn from the entirety of the WDI, which contains 1,265 variables in total. We trim this larger set of variables in the following manner: First, we drop any variable that is not reported by at least one country in every year from 1980-2010. In so-doing, we ensure that our index has constant meaning throughout the time period we cover.

Second, we exclude all variables that are – or appear to be – measured for only a subset of countries. For instance, we exclude Official Development Assistance (ODA) measures and variables related to holdings of external debt, since these are only routinely reported for – respectively – low income and highly indebted nations. The data generating process underlying the reporting of these variables systematically differs from that for other data in the WDI, and thus including these variables in our model would be inappropriate.

Third, we exclude any data that is derived from World Bank sponsored surveys (e.g., the Doing Business surveys) or that is an index compiled by external entities. We also exclude (with one exception, detailed below), variables that are derived from World Bank calculations – e.g., growth rates and purchasing power parity (PPP) conversion factors. Our index is intended to reflect *governments'* disclosure of data, and we thus attempt to confine our data to variables reported with input from governments.

Fourth, we drop variables that are linear combinations of other variables. For instance, data are reported on the total labor market participation rate, the male labor market participation rate, and the female

[8]A complete listing of the variables included in our analysis is presented in the Appendix.

[9]In settings where a researcher does wish to impose strong theoretical priors, items can be excluded. See, for example, Gates et al. (2006).

labor market participation rate – as well as relevant population figures. Of the three labor market participation measures, we include only the female participation rate.

Fifth, we delete multiple references to the same underlying measure that are reported in different units. For instance, national accounts data are typically reported in current US dollars, current local currency units (LCU), constant US dollars, constant LCUs, and purchasing power parity US dollars. When measures are reported in this manner, our first preference is to keep only the variable reported in constant US dollars. If no such variable is available, we then keep only the variable reported in constant LCU. In all instances, we keep only one such reference. Our preference for reporting in constant US dollars would seem to contradict our third requirement for inclusion. After all, most governments report values in their local currency, conversions are subsequently conducted by the World Bank. We prefer the constant US dollar values, however, as these values incur an additional layer of World Bank scrutiny. The World Bank notes that constant dollar values may be left as missing when "originally reported growth rates appear to be unrealistic."[10]

We also restrict the set of countries included in our analysis. We first exclude any country that did not exist for the entirety of the 1980-2010 period.[11] We do this for two reasons: First, the WDI only includes countries currently in existence. For countries that came into existence after the WDI begins its coverage, values for dates prior to independence are constructed – wherever possible – based on existing data (e.g., data from prior rulers and/or information made available by the current leadership based on pre-independence records). Second, rectangularizing the dataset eases the construction of our algorithm. For similar reasons, we exclude modern countries that are formed by the union of preexisting states during the 1980-2010 period – i.e., Germany and Yemen. Finally, we drop all micro-states from the dataset – that is, we only include states that maintained a population of 500,000 or more throughout the 1980-2010 period. After dropping states that do not meet these criteria, we are left with 125 countries, each of which is observed for 31 years, for a total of 3,875 observations.

In our model, we let $y_{j,c,t} \in \{0, 1\}$ denote an indicator equal to 1 if country c reports WDI variable j in

[10] `http://data.worldbank.org/about/faq/specific-data-series`, last accessed March 5, 2013.

[11] To determine dates of independence, we rely on the DD dataset of Cheibub, Gandhi and Vreeland (2010) and the PWT 7.1 (Heston, Summers and Aten, 2012). We merge the WDI data with these two datasets and drop all observations from the WDI that are not included in both. We do, however, keep countries that fractured in the index, if the largest constituent portion of that country remained independent. So, Russia is treated as the continuation of the USSR, Ethiopia remains in the dataset after the succession of Eritrea, Indonesia remains after the succession of East Timor, etc.

year t and equal to 0 otherwise. We then estimate

$$Pr(y_{j,c,t} = 1 | transparency_{c,t}) = logit(\delta_j + \beta_j transparency_{c,t}) \tag{1}$$

where δ_j is the difficulty parameter and β_j the discrimination parameter for item j. The term $transparency_{c,t}$ is the measure of a given country-year's propensity to disclose data, which is to be estimated. The $logit$ function is a logistic transformation. Thus, changes in δ_j reflect the degree to which countries, on average, report a given variable drawn from the WDI. Changes to β_j reflect the degree to which the outcome of one item predicts the outcome of other items (Gelman and Hill, 2006). Simply taking the fraction of reported variables as an indicator of the transparency of each country would therefore be approximately equivalent to assuming δ_j and β_j are constant across all variables j reported to the WDI. Our approach here is thus far more general than simple averaging, and lets the data reveal the amount of information contained in each item (WDI variable).

In our model there are 240 items j, 125 countries c and 31 years t. We thus estimate a system of 240 equations (all of the form of equation 1) with 3,875 observations.

The model represented by equation 1 is only identified up to an affine transformation. To fix the scale and location of the index, we recenter simulated index values – by subtracting the mean and dividing by the standard deviation – in the year 1980 at each iteration of the MCMC procedure (Knorr-Held, 1999). Initial values for the index in year 1980, the first year of the dataset, are drawn from a diffuse normal prior $Transparency_{i,1980} \sim N(0, 100)$ prior to recentering.[12] We fix the direction of the index by constraining Cuba to have negative transparency scores and Sweden to have positive transparency scores, using half-normal priors for each.[13]

In this instance, we encounter an additional modeling difficulty. Since our data are time-series cross-sectional, we cannot treat each observation of $transparency_{c,t}$ as independent of the observation $transparency_{c,t-1}$. We therefore employ a system of random-walk priors for each value of $transparency_{c,t}$ after 1980. Our system of priors is thus: $transparency_{c,t} \sim N(transparency_{c,t-1}, \frac{1}{\tau_c}) \; \forall \; t > 1$ and $transparency_{c,t} \sim N(0, 1)$ for $t = 1$. τ_c thus acts as a country specific smoothing parameter, since the degree to which an

[12] We would like to thank an anonymous reviewer for suggesting this approach.

[13] We choose Cuba and Sweden as anchors because, in previous runs of the model without these constraints, Cuba consistently as one of the lowest and Sweden as one of the highest countries on our index.

estimated value of $transparency_{c,t}$ shrinks back towards the prior mean is inversely proportional to the variance of the prior distribution (Jackman, 2009). Similar priors are often used in the ideal point estimation literature when working with dynamic data (see, for instance Martin and Quinn, 2002). We estimate the parameter τ_c for each country in our data, and give this term a prior $\tau_c \sim Gamma(20, 0.25)$.[14]

We place diffuse normal priors on the difficulty and discrimination parameters δ_j and β_j, such that

$$\begin{pmatrix} \delta_j \\ \beta_j \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}).$$

We estimate this model employing a Markov Chain Monte Carlo algorithm run from JAGS 3.3.0. The model is estimated using 2 chains of 10,000 iterations each, the first 5,000 of which serve as a burn-in period. Gelman-Rubin convergence diagnostics reveal that the model converges. Of the 4,480 parameters in the model, the highest valued Gelman-Rubin statistic is 1.079. We report convergence statistics for all variables in the Appendix.

To assess model fit, we rely on posterior predictive checks (Gelman et al., 2000). We sample repeatedly from the posterior distributions of all $transparency_{c,t}$, difficulty (δ_j) and discrimination (β_j) terms to generate the predicted probability a given country c reports a given variable j in a given year t, which we denote $p_{j,c,t}$. We calculate the deviation between these predicted probabilities and the observed data and express this deviation in a single measure. Specifically, we construct a measure analogous to Herron's (1999) expected percentage correctly predicted (ePCP), which we term the total expected percentage correctly predicted (TePCP). The TePCP is defined as follows:

$$TePCP = \frac{1}{JCT} \sum_{j=1}^{J} \sum_{c=1}^{C} \sum_{t=1}^{T} [y_{j,c,t} p_{j,c,t} + (1 - y_{j,c,t})(1 - p_{j,c,t})]$$

where $J = 240$ denotes the total number of items in the model, $C = 125$ denotes the total number of countries, and $T = 31$ denotes the total number of years. As noted above $y_{j,c,t} \in \{0, 1\}$ denotes whether variable j is reported by country c in year t. The TePCP thus takes on a value in the open unit interval, and a notional value of 1 would indicate perfect predictive power. We run 500 simulations, sampling repeatedly

[14] A Gamma distribution with a shape parameter 20 and scale parameter 0.25 has a mean of 5 and a variance of 1.25.

from the posterior densities of each model parameter, to generate 500 TePCP values. These values reveal that the model fits the data well: the average TePCP score is 0.796, and 95 percent of all TePCP scores fall between 0.795 and 0.796.[15]

4 The Value-Added of Item Response

Previous studies (Hollyer, Rosendorff and Vreeland, 2011; Williams, 2009) compute analogous measures of data disclosure based on the fraction of variables reported to the WDI or other international datasets. Given the computational intensity of the IRT approach, it is reasonable to ask whether there is any value-added in fitting an IRT model to these data rather than relying on the fraction of non-missing observations. As we demonstrate below, the IRT approach does offer a more nuanced view of transparency than does the simple sum-score measure. The IRT model particularly improves our ability to distinguish among extreme types – i.e., highly transparent and opaque states – relative to the sum-score index, which clusters these observations together.[16]

In theory, the IRT model offers a substantial advantage over simply computing the fraction of variables reported to international agencies. By construction, the latter approach assumes that the reporting of any one variable is equivalent to the reporting of any other. This assumption is highly restrictive. The IRT approach, by contrast, allows each variable to contribute differentially to index values – the contribution of each variable to the index will depend on the extent to which the reporting of this variable is correlated with the reporting of other variables. The model adjusts for the fact that some variables are simply much more commonly reported than others, and for the fact that the reporting of certain terms is more predictive of a general tendency to disclose than the reporting of other terms.

Nonetheless, it is possible that, in practice, each item contained in the WDI is as informative as any other. If this is the case, the IRT approach would be roughly equivalent to simply computing the fraction

[15] By way of comparison, we fit a model consisting only of constant terms – i.e., the difficulty parameters δ_j. This is roughly equivalent to assuming that each observation of $p_{i,j}$ is equal to the empirical frequency with which a given variable j is reported in the WDI data. This process produces a mean TePCP of .707 with 95 percent of all observations falling between .706 and .708. Thus, the IRT model results in a roughly 30% marginal reduction in error. We additionally explore whether a two dimensional IRT better fits the data. The two dimensional model only modestly improves model fit, producing an average TePCP of .8381, with 95 percent of all values falling between .8378 and .8383. We present a graphical comparison of the simulated TePCP values from each model in the Appendix.

[16] In the Appendix, we explicitly compare the use of the a sum-score and the IRT measure in an applied setting.

of variables reported. It might then be advisable to rely on the simpler and less computationally intensive approach rather than going through the trouble of fitting an item response model.

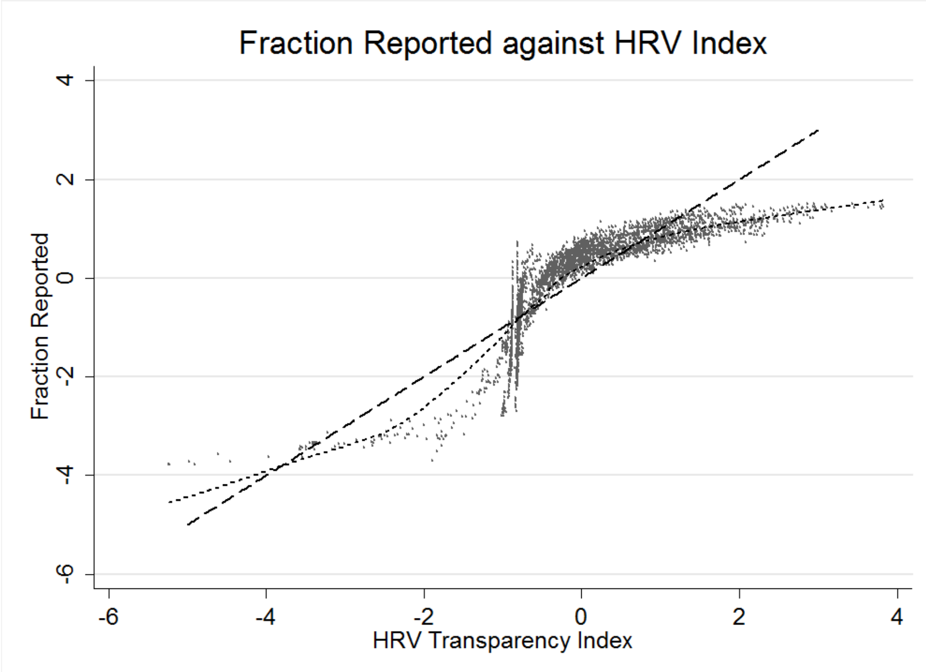
We can assess this empirical possibility in two ways: First, we can simply compare the HRV index values to those based on the fraction of the 240 variables in our model that are non-missing. Second, we can examine the estimates of the difficulty and discrimination parameters (δ_j and β_j) from the item response model. If these values differ significantly across items, then the IRT is a more appropriate model for these data than the raw frequency of reporting.

Figure 2 contains a scatterplot, together with a lowess regression line, of mean HRV index values against the fraction of the 240 WDI variables used in our analysis reported by a given country in a given year. We have standardized both variables by subtracting their means and dividing by their standard deviations. It is immediately apparent from the figure that these two measures are related – the raw correlation between the mean HRV index score and the fraction of variables reported is 0.84. However, the scatterplot in Figure 2 departs substantially from the dashed 45 degree line that represents a perfect correlation. Notably, the slope of the fitted lowess curve is substantially less than one for nearly all values of the HRV index. This implies that the HRV index is more sensitive to variations in reporting than is the fraction of variables reported – high-scoring countries do better (relative to the mean) on the HRV index than on the fraction of variables reported, low-scoring countries do worse. The relationship between the two-terms is also non-linear – the fitted lowess line is sigmoid shaped. HRV thus detects greater variation among countries receiving relatively extreme scores than does the fraction reported measure; and finds less variation for countries with middling (slightly below average) scores.

Another means by which one can assess the value-added of the IRT approach is via the examination of the difficulty (δ_j) and discrimination (β_j) parameters for the 240 items in our model. Recall that the use of the fraction of variables reported as a measure of data dissemination is equivalent to holding these parameters constant across all variables examined. If the estimates of these parameters from our model vary little across items, one might therefore conclude that the IRT model is unnecessary.

In fact, the estimates of these parameters vary considerably. Point estimates for the difficulty parameters vary across items from a minimum of -4.6 to a maximum of 16.5 (with a standard deviation of 4.18). Point estimates of the discrimination parameters vary from a low of -0.14 to a high of 58 (with a standard

Figure 2: Fraction Reported against Mean HRV Index Scores



The fraction of the 240 variables used to construct the HRV Index reported in a given country year plotted against HRV Index scores. Both measures have been standardized by subtracting the mean and dividing by the standard deviation. A 45 degree dashed line runs through the origin. The dotted line depicts the lowess curve fitted to these values.

deviation of 14.3). Figures 3 and 4 plot these estimates, and their 95 percent highest posterior density intervals, for the 25 most and 25 least discriminating items in our data.[17]

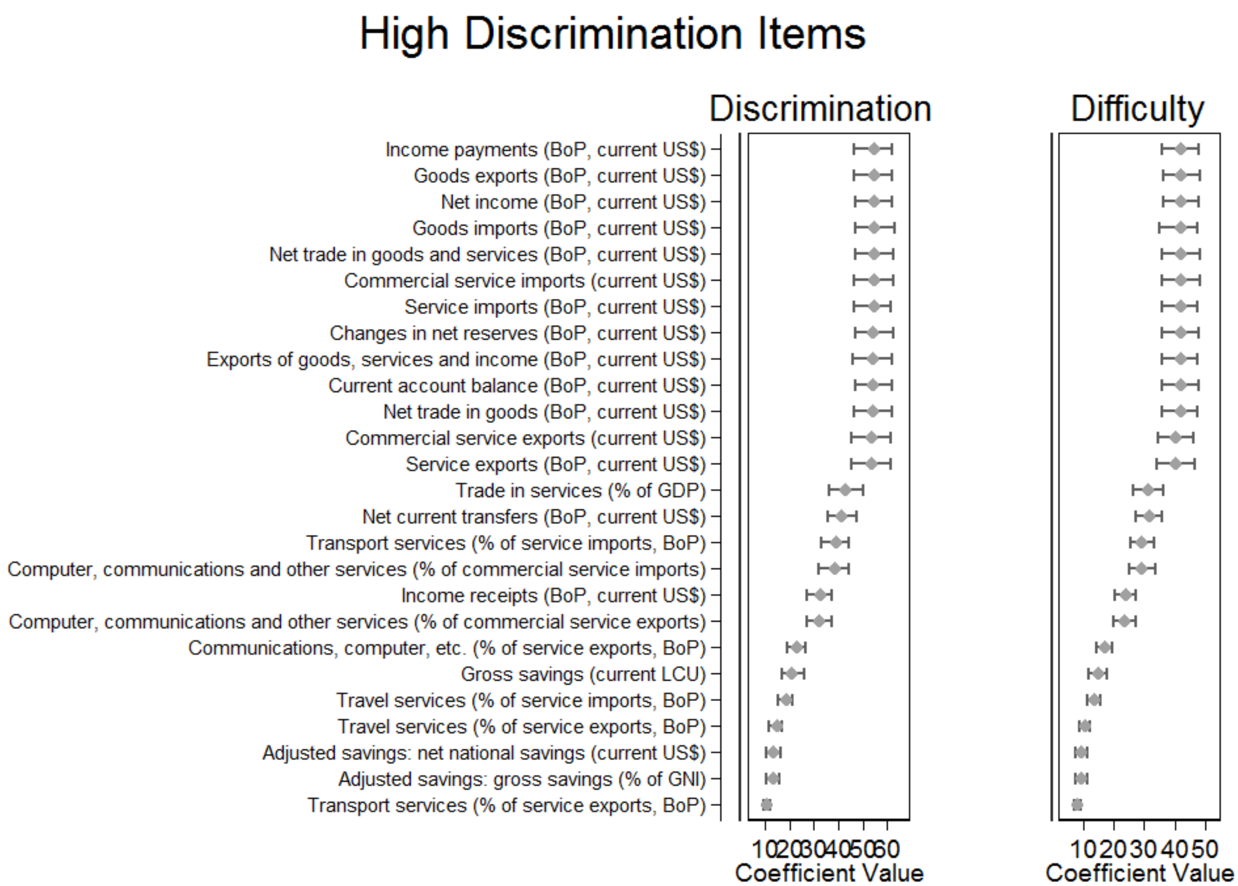
The examination of these parameters also provides some insight into the validity of our index. If our measure of transparency is to be relevant to political processes, it must reflect the disclosure of *politically relevant* data. If the most discriminating items are systematically obscure politically irrelevant variables, one might conclude that our index reflects bureaucratic technocracy, rather than public data disclosure.

The most discriminating items overwhelmingly relate to trade and investment. These include measures of the current account balance, goods and services exports and imports, and changes in reserves. These items tend also to have positive and large difficulty parameters, implying that they are reported with some frequency. By contrast, the least discriminating parameters consist of population measures, measures of

[17] Analogous plots for all items are presented in the Appendix.

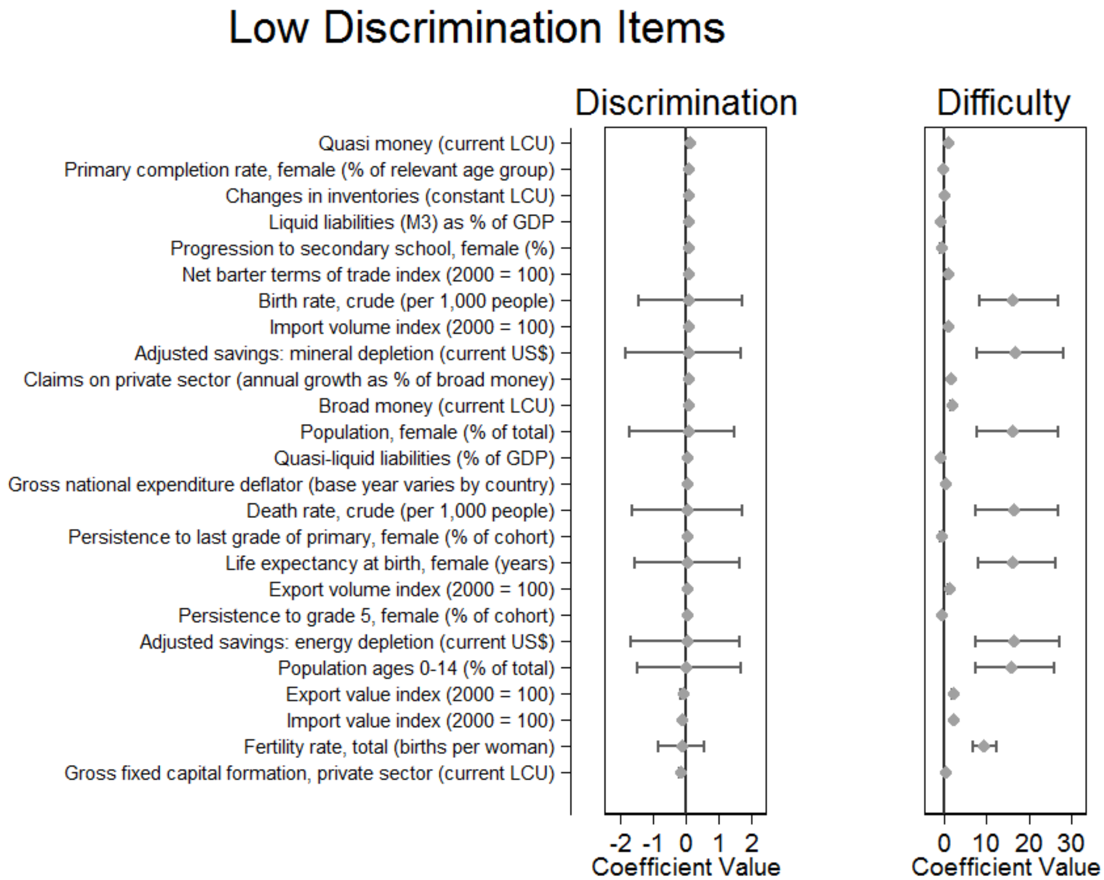
the money supply, and a variety of detailed measures related to education (e.g., the pupil-to-teacher ratio). The highly discriminating items are thus composed of measures with clear political relevance. A handful of the less discriminating items (e.g., the education items) may be politically relevant in some settings. But, the majority of these items (fertility rates, population values, values of M3 as a percentage of GDP) are unlikely to be at the center of political concern. We therefore conclude that our index reflects the disclosure of politically relevant data and not merely technocracy.

Figure 3: Parameter Estimates, Highest Discrimination Scores



Discrimination and difficulty parameters, and 95 percent highest posterior density intervals, for the 25 items with the highest discrimination parameter scores. Parameter values are plotted on the x-axis and mean estimated values are denoted by diamonds. 95 percent highest posterior density intervals are denoted by whiskers. Item names are noted on the y-axis

Figure 4: Parameter Estimates, Lowest Discrimination Scores

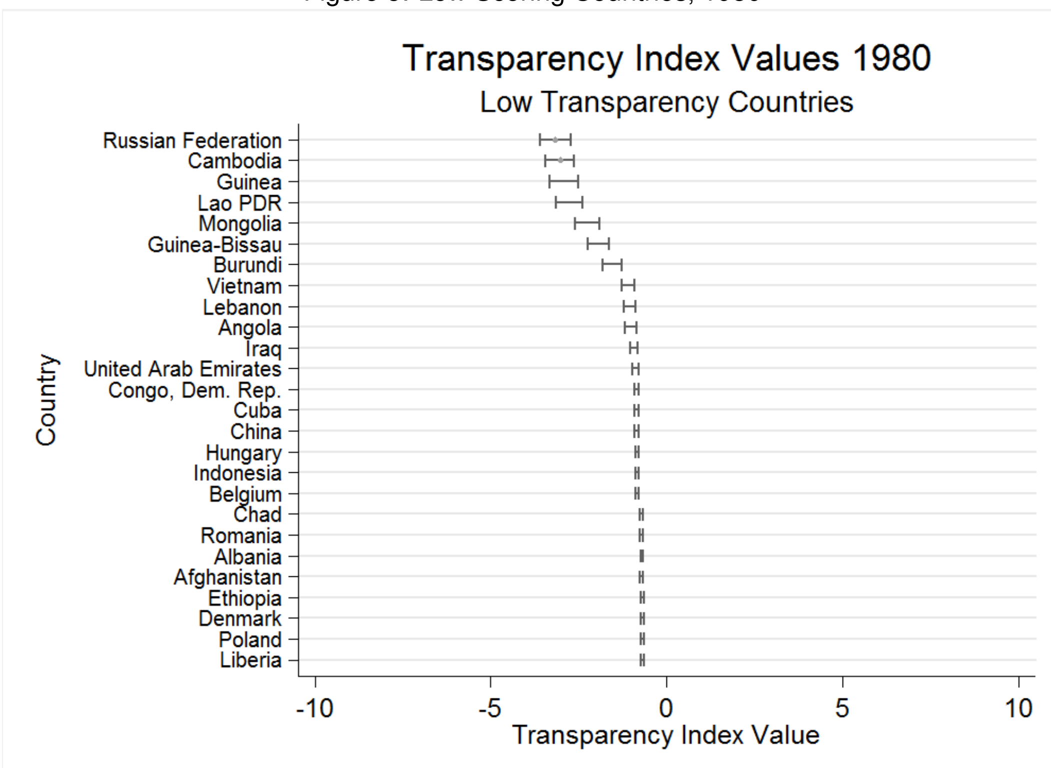


Discrimination and difficulty parameters, and 95 percent highest posterior density intervals, for the 25 items with the lowest discrimination parameter scores. Parameter values are plotted on the x-axis and mean estimated values are denoted by diamonds. 95 percent highest posterior density intervals are denoted by whiskers. Item names are noted on the y-axis

5 Face Validity

To assess the validity of our estimates, we first examine the estimated values of the transparency index in the 1980 cross-section. We report estimated index values and 95 percent highest posterior density intervals for the 25 lowest scoring countries in Figure 5 and for the 25 highest scoring countries in Figure 6. As can be seen from the estimates, nearly all the highest scoring countries are members of the OECD. While the lowest scoring countries consist of Russia (the former Soviet Union), Cambodia, Guinea, Laos, Mongolia and Guinea-Bissau – all states we would expect to score poorly on transparency during this period. We can also see that the highest scoring country in this period (Spain) differs significantly in its score from other high-scoring countries such as the UK. This differentiation is even more dramatic among low scoring countries. The Soviet Union is significantly less transparent than Guinea which is, in turn, significantly less transparent than Mongolia. This provides evidence that not only do our index values coincide with common notions of transparency, but we are also able to discriminate among high and low scoring countries. We are able to differentiate between countries not only at the extremes – as plotted in Figures 5 and 6 – but also in the middle range of transparency scores.

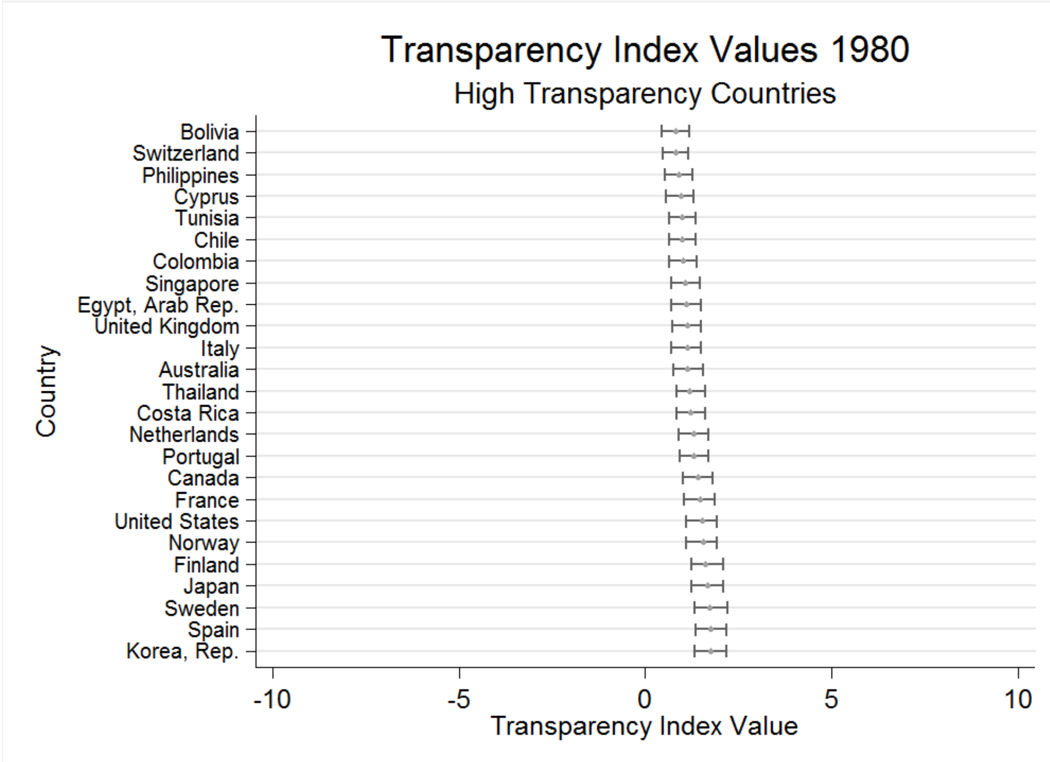
Figure 5: Low Scoring Countries, 1980



A cross section of the 25 lowest scoring countries on the HRV Index from 1980. Three letter ISO country codes are plotted on the y-axis. Index scores are plotted on the x-axis. Mean predicted transparency index values are indicated by diamonds, while the whiskers denote 95 percent highest posterior density intervals. Countries are plotted in ascending order of mean HRV Index score.

The general correspondence between HRV scores and common preconceptions of countries' levels of transparency does not imply, however, that the examination of these data does not yield a few surprises.

Figure 6: High Scoring Countries, 1980



A cross section of the 25 highest scoring countries on the HRV Index from 1980. Three letter ISO country codes are plotted on the y-axis. Index scores are plotted on the x-axis. Mean predicted transparency index values are indicated by diamonds, while the whiskers denote 95 percent highest posterior density intervals. Countries are plotted in ascending order of mean HRV Index score.

Figure 5 reveals, for instance, that several first world countries – particularly Belgium and Denmark – score remarkably poorly on the HRV index. Indeed, both countries experience substantial volatility in HRV scores over time. Denmark exhibits remarkable variation in the frequency of data disclosure over time – whether this is measured in terms of HRV scores or of the fraction of variables that are non-missing. It regularly oscillates between periods of high and low disclosure, with particularly large changes in the mid-1980s and mid-2000s. Belgium, on the other hand, exhibits consistently low levels of disclosure until 2002, when reporting jumps dramatically. This change is a particularly extreme example of volatility in European HRV scores particularly – among established EU members – around the time of the introduction of the euro, and – among new EU members – around the time of accession. Such volatility is likely due EU (and euro-area) reporting requirements.

In addition to comparing across countries cross-sectionally, we examine longitudinal variation in our estimates to assess their validity. Based on anecdotal accounts, the level of transparency in many countries in our sample should be shifting over time. We present our estimates of the level of transparency in several such countries below.

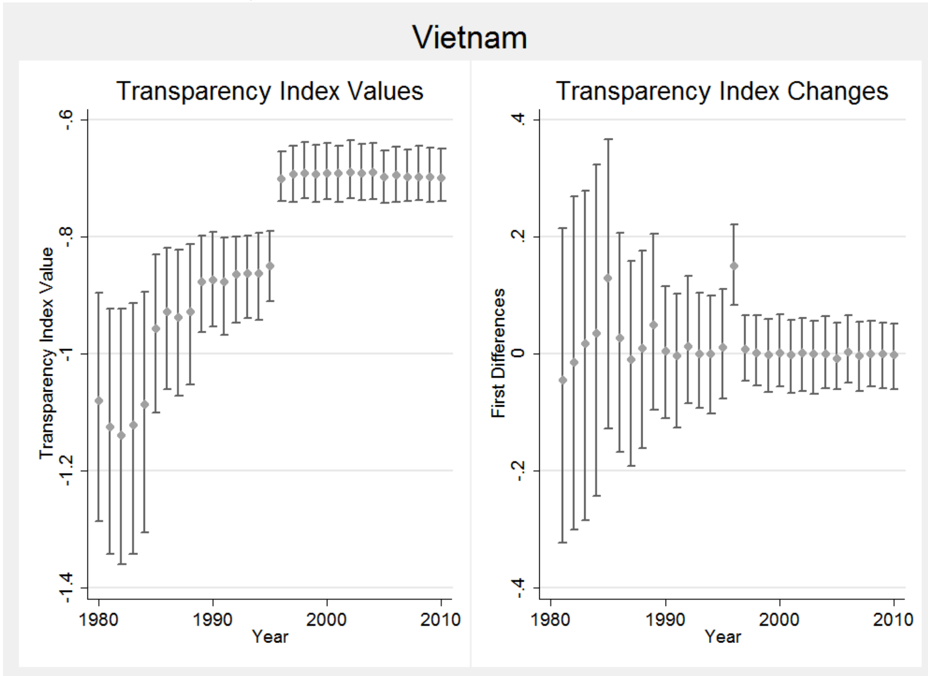
For instance, Vietnam experienced enormous economic and political changes over the 1980-2007 period. These changes included economic liberalization, the creation of (relatively) more representative institutions, and the establishment of diplomatic relations with many nations. The 1986 Party Congress is widely seen as ushering in the period of liberalization (Fforde and De Vylder, 1996; Riedel and Turley, 1999). The process of economic and political liberalization continued throughout the 1990s and early 2000s (Malesky, 2008, 2009; Gehlbach and Malesky, forthcoming). One might therefore expect that Vietnam would grow substantially more transparent over time, particularly after 1986. Figure 7, which plots our predicted transparency index scores and 95 percent highest posterior density intervals for Vietnam over time, shows that our index nicely reflects this history.

From the 1970s through the 1980s, Tanzania was a single party state run by the Chama Cha Mapinduzi (CCM) party. The CCM advocated the socialist principles and import-substituting growth strategy espoused by the founding father of the nation, Julius Nyerere. In the 1980s, this system came under increasing strain due to chronic current account imbalances, IMF pressure, and the international collapse of communist states (Kiondo, 1992; Mmuya and Chaligha, 1992, 1994; Vreeland, 2003). In 1992, the CCM agreed to open the political system to multiparty elections. Tanzania also embarked on a system of significant economic liberalization during the early 1990s (Hoffman, 2011).

Figure 8 plots our predicted transparency scores and 95 percent highest posterior density intervals for Tanzania over time. An abrupt and significant upwards shift in transparency values is noticeable in 1990, followed by further improvements throughout the early 1990s. These shifts coincide nearly perfectly with the period of economic and political liberalization.

Other countries exhibit sharp declines in estimated levels of transparency over time. A stark example is visible in Figure 9, which plots Somalia's HRV index scores over time. Throughout the 1970s and 1980s, Somalia was ruled by the nominally communist government of Mohamed Siad Barre. Following a military defeat to Ethiopian forces in 1978, the Siad Barre government maintained control by relying on a system

Figure 7: Vietnam HRV Scores Over Time



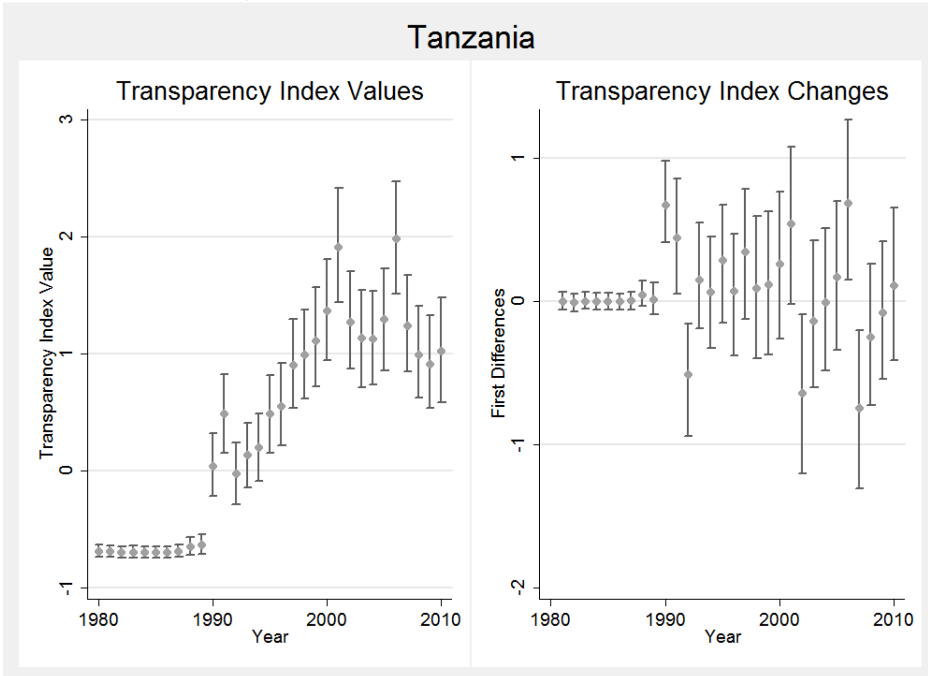
A longitudinal plot of Vietnam's HRV Index scores from 1980-2010. HRV Index scores are on the y-axis, while time is measured on the x-axis. Diamonds denote mean predicted index scores. Whiskers denote 95 percent highest posterior density intervals.

of clan-based patronage.[18] This government is associated with moderate to low HRV Index scores of between 0 and -0.15. By the late 1980s, this government faced increasing armed domestic opposition and infighting between clans once supportive of the government. Unrest grew such that, in 1991, Siad Barre was forced to flee Mogadishu in a tank, taking with him the foreign exchange reserves of the central bank (Prunier, 1996, 45). The collapse of the government ushered in a period of civil war during which Somalia became the paradigmatic example of a failed state. The collapse of the Siad Barre government coincides perfectly with a dramatic – and statistically significant – fall in Somalia's HRV index score. In 1990, Somalia's predicted HRV score fell from -0.75 to -2.89. The following year saw another dramatic drop to -4.92, followed by further declines. Somalia's transparency score stabilizes around values of -7, among the lowest in our sample.

Nor are the HRV time series merely consistent with montonic shifts in transparency over time. For

[18] For an overview of the rise and fall of the Siad Barre regime, see Prunier (1996).

Figure 8: Tanzania HRV Scores Over Time



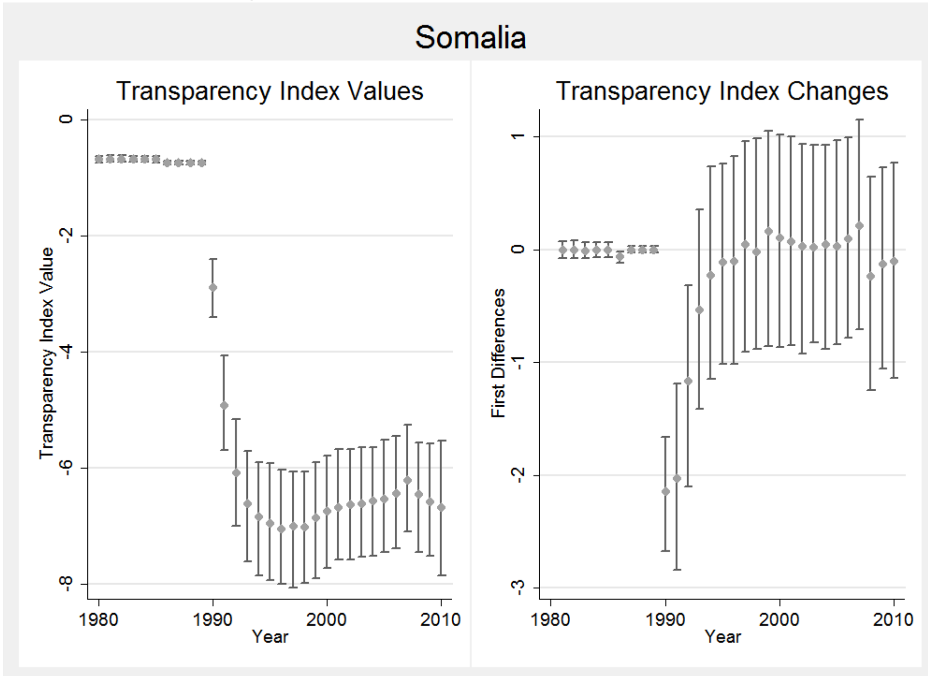
A longitudinal plot of Tanzania's HRV Index scores from 1980-2010. HRV Index scores are on the y-axis, while time is measured on the x-axis. Diamonds denote mean predicted index scores. Whiskers denote 95 percent highest posterior density intervals.

instance, Argentina experienced massive swings in the level of economic liberalization between 1990 and 2010. In the early 1990s, then Menem government introduced policies to curtail the hyperinflation that had run rampant throughout the 1990s (Treisman, 2004). The policies implemented by Menem remained largely in effect until the Argentinian debt crisis and the subsequent election of Néstor Kirchner in 2003. Particularly relevant for data dissemination: Kirchner's successor and wife – Cristina Fernández – seized control of the (previously apolitical) statistics institute in 2007. Since that time, inflation statistics have been systematically doctored – largely to reduce government payments on inflation-linked debt – leading the IMF to censure Argentina non-compliance with the Fund's rules governing the reporting of statistics in February of 2013.[19]

Plots of the Argentina's HRV scores over time are presented in Figure 10. Transparency levels in

[19] *The Economist,* "Motion of Censure." February 9, 2013. http://www.economist.com/news/americas/21571434-fund-blows-whistle-motion-censure.

Figure 9: Somalia HRV Scores Over Time



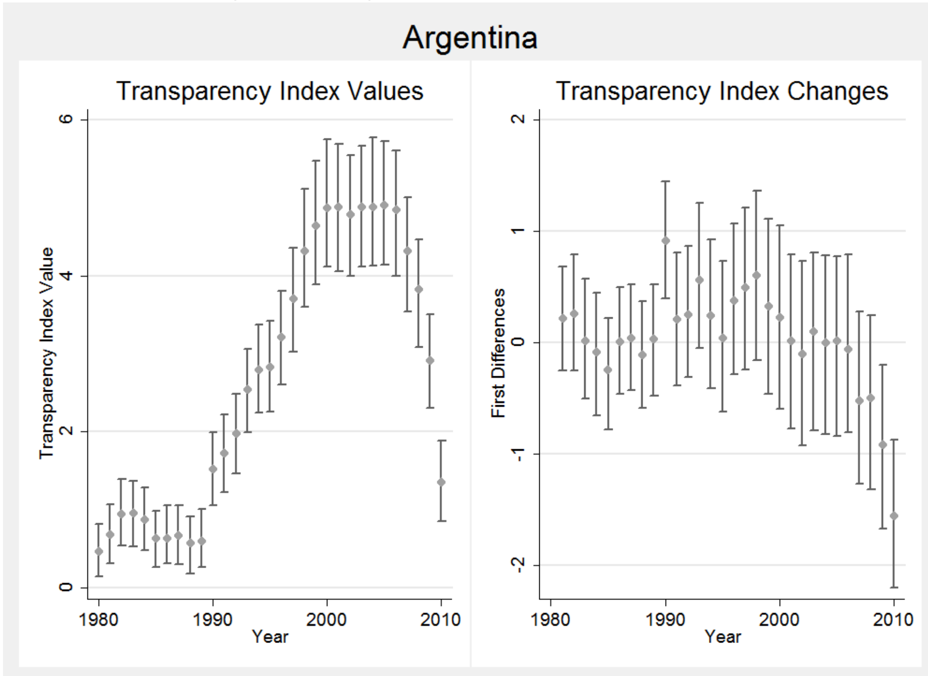
A longitudinal plot of Somalia's HRV Index scores from 1980-2010. HRV Index scores are on the y-axis, while time is measured on the x-axis. Diamonds denote mean predicted index scores. Whiskers denote 95 percent highest posterior density intervals.

Argentina exhibit a clear non-monotonicity. Scores begin a dramatic increase in 1991 – the year in which the Convertability Plan was introduced – and these gains continue throughout the 1990s. Transparency plateaus during the debt crisis, before declining sharply in the late 2000s. The first year of this decline is 2007 – the same year in which Fernández took control of the statistics institute.

6 Validation Exercise: Transparency and Governance

Political scientists have long been concerned with the importance of political accountability and its effects on government performance. A seminal contribution to this literature is that of Adserà, Boix and Payne (2003) (henceforth, ABP), which examines the relationship between democracy and the free flow of information and their interaction and three forms of governance: (1) bureaucratic quality, (2) the rule of law, and (3) corruption. ABP also examine the relationship between these terms and the risk of expropriation, which

Figure 10: Argentina HRV Scores Over Time



A longitudinal plot of Argentina's HRV Index scores from 1980-2010. HRV Index scores are on the y-axis, while time is measured on the x-axis. Diamonds denote mean predicted index scores. Whiskers denote 95 percent highest posterior density intervals.

we do not replicate here for reasons of data availability. Democracies, they find, perform significantly better than autocracies in all four measures – and the role of democracy is further accentuated when information (proxied by the level of circulation of daily newspapers) flows freely.

We replicate the work of ABP below, using both the HRV Index of data dissemination and the circulation of daily newspapers as alternative measures of transparency.[20] If each is associated with governance

[20]Ours is not an exact replication of ABP, as several of the datasets they use have been updated and modified. Where they use the ICRG's indexes of corruption, the rule of law, bureaucratic quality, and the risk of expropriation, we use the most recent ICRG measures of corruption, law and order, and bureaucratic quality. ABP examine the risk of expropriation, while we do not, because the current ICRG does not include such a measure. Whereas the corruption, rule of law, and bureaucratic quality indexes in ABP's specifications all ranged from 0-6; the bureaucratic quality measure in our specifications ranges from 0-4. ABP measure democracy using the Polity III democracy measure (rescaled to lie between 0 and 1). We rely on the more commonly used polity2 measure drawn from the Polity IV dataset. Finally, ABP rely on the daily newspaper circulation per 1,000 individuals measure drawn from the year 2000 WDI as their measure of transparency. We append additional observations of this measure (up to 2004) from the 2011 WDI. Our specification also differs slightly from that of ABP. They include an interaction between newspaper circulation and democracy in their model, but assume that newspaper circulation has no direct effect on transparency. We prefer not to make such an assumption and, following the advice of Brambor, Clark and Golder (2005), include the constitutive terms of all interactions in our specifications.

outcomes – even after controlling for the other – these results support the contention that different dimensions of transparency have distinct implications for government accountability (we develop this argument in Hollyer, Rosendorff and Vreeland, 2013*b*).[21]

We follow ABP in treating each governance indicator as a linear function of democracy (here proxied by the polity2 score), GDP *per capita* in 2005 constant purchasing power parity US dollars, transparency, the interaction of transparency and democracy, and a lagged dependent variable. The model is thus

$$\begin{aligned} governance_{c,t} = \gamma_1 + \gamma_2 governance_{c,t-1} + \gamma_3 democracy_{c,t} + \gamma_4 transparency_{c,t} + \\ \gamma_5 democracy_{c,t} * transparency_{c,t} + \gamma_6 newscirc_{c,t} + \gamma_7 newscirc_{c,t} \times democracy_{c,t} + \\ \gamma_8 GDP_{c,t} + \epsilon_{c,t}. \end{aligned} \tag{2}$$

(Note that the presence of a lagged dependent variable makes the model dynamic. Thus, the steady state association between these terms and governance is given by dividing each coefficient value by $1 - \gamma_2$.)

Also following ABP, we collapse yearly observations into half-decade intervals, taking the mean of all terms except newspaper circulation in each five year period. As newspaper circulation is only measured in 5 year increments prior to 1997, we simply take the newspaper circulation figures from the first year in each period as the observation. Collapsing the data helps to adjust for slow movement in the governance indicators over time and helps to reduce measurement error in these terms.

Since $transparency_{c,t}$ is measured with error, estimating the above equation via OLS would bias coefficient values. To avoid this problem, we estimate the above equation via MCMC, nesting the regression equation and the measurement model in the same MCMC algorithm (Fox and Glas, 2003). Since the ICRG measures used as outcomes in the regression model are not available for all countries and years in the HRV index, we estimate a measurement model identical to the baseline model described above for the 106 countries contained in both datasets, for the years 1986-2008. Within the same algorithm, we collapse the resultant $transparency$ scores into their five year averages and estimate equation 2. We place diffuse

[21] We do not intend for these results to be interpreted causally. As noted above, our transparency scores are likely to be correlated with state capacity, which may act as an omitted variable driving the relationship with governance outcomes. This problem is likely to be particularly severe with respect to measures of bureaucratic quality. While controlling for GDP *per capita* helps to adjust for this problem, it may not solve it. We should note, that such bias is unlikely to explain why the HRV index is more strongly associated with governance outcomes in autocratic – as opposed to democratic – states. Our purpose here is to ascertain the validity of our measure by replicating a seminal work in the literature.

priors on the coefficients γ such that:[22]

$$\gamma \sim N(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1,000 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1,000 \end{pmatrix}).$$

We run this algorithm in JAGS 3.3.0, using 2 chains of 5,000 iterations each, the first 3,000 of which serve as a burn-in period. Convergence statistics on all parameters indicate that the model converged – though, with 2,734 parameters, a few parameters indicate high Gelman-Rubin diagnostic scores by chance (Gill, 2008). All but 5.8 percent of parameters have Gelman-Rubin diagnostics of less than 1.1.[23]

We report coefficient estimates, and 95 percent highest posterior densities, in Table 1. Table 1 reveals two distinct patterns. First, broadly speaking, it confirms the main finding of ABP: newspaper circulation is associated with improved governance only under democracy, where the press is presumably free. Secondly, data dissemination is associated with improved governance in non-democracies – its correlation with governance in the most democratic states is weak at best. This finding is consistent, with the theoretical notion that data dissemination may be important for forms of accountability involving collective action, but is less critical where electoral forms of sanctioning political leaders are available (see Hollyer, Rosendorff and Vreeland, 2013*a*).

To ease in the interpretation of the results of these regressions, we plot the marginal contemporaneous effect of a one-standard deviation change in each both the HRV score and in newspaper circulation, for each possible value of polity, in Figures 11, 12, and 13 (Brambor, Clark and Golder, 2005). In each graph, the values of the polity score are plotted on the x-axis, and the values of the marginal effect are plotted on the y-axis. 95 percent highest posterior density intervals are denoted by dashed lines. Plots for the HRV index and for newspaper circulation are plotted side-by-side on a common y-axis, for ease of comparison.

The HRV index is more strongly correlated with bureaucratic quality and law & order than is newspaper circulation, in all but the most democratic states. Neither measure of transparency is highly correlated with

[22]We have conducted robustness checks with normal priors with variances of 10,000 and 500. Results available from the authors on request.

[23]Results available from the authors on request.

Table 1: Governance Measures Regressed on Transparency

	Bur. Quality	**Law & Order**	**Corruption**
Lag Dep. Variable	0.62	0.55	0.59
	[0.55, 0.67]	[0.47, 0.64]	[0.51, 0.68]
Polity2	8×10^{-3}	-0.03	0.01
	[-4×10^{-3}, 0.02]	[-0.05, -0.01]	[-7×10^{-3}, 0.03]
HRV Index	0.10	0.15	0.06
	[0.04, 0.15]	[0.06, 0.23]	[-0.01, 0.14]
Polity2 $\times$	-7×10^{-3}	-0.02	-0.01
HRV Index	[-0.01, 4×10^{-4}]	[-0.03, -5×10^{-3}]	[-0.02, -2×10^{-3}]
Newspaper Circ.	1×10^{-3}	3×10^{-4}	1×10^{-3}
	[-5×10^{-4}, 3×10^{-3}]	[-2×10^{-3}, 3×10^{-3}]	[-1×10^{-3}, 4×10^{-3}]
Polity2 $\times$	-3×10^{-5}	2×10^{-4}	1×10^{-4}
Newspaper Circ.	[-2×10^{-4}, 1×10^{-4}]	[-6×10^{-5}, 4×10^{-4}]	[-1×10^{-4}, 3×10^{-4}]
GDP *per capita*	2×10^{-5}	3×10^{-5}	8×10^{-6}
	[7×10^{-6}, 3×10^{-5}]	[7×10^{-6}, 4×10^{-5}]	[-5×10^{-6}, 2×10^{-5}]
σ	0.47	0.73	0.70
	[0.43, 0.51]	[0.66, 0.79]	[0.64, 0.76]
N	261	261	261

Results of a regression of ICRG governance indicators against differing measures of transparency. 95 percent highest posterior density intervals are presented in brackets.

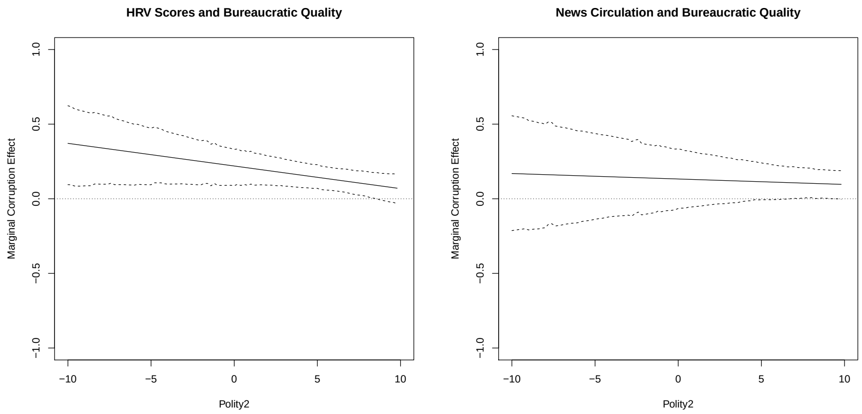
the level of corruption – the HRV index outperforms newspaper circulation in non-democracies, newspaper circulation is a better predictor of corruption in democracies.

Our results thus demonstrate that the HRV index is robustly correlated with governance outcomes, controlling for an alternative form of transparency (i.e., newspaper circulation). This correlation is broadly stronger than that between newspaper circulation and either bureaucratic quality or law and order. And data dissemination is strongly linked to the quality of governance in autocracies.

7 Conclusion

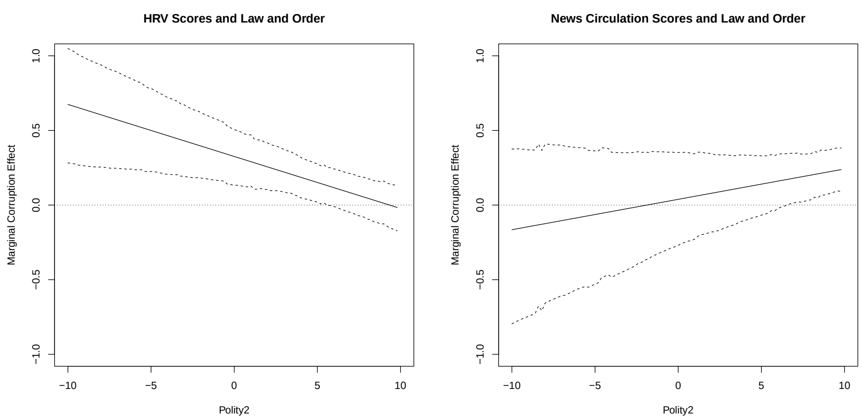
We provide an index measuring data dissemination with many desirable properties. Our index is based on objective criteria, rather than subjective judgments and reflects the dissemination of credible information. Moreover, unlike many cross-national indexes, the HRV index values are unique up to an affine transfor-

Figure 11: Marginal Effects of a Standard Deviation Change in Transparency on Bureaucratic Quality



Plots of the marginal effect of transparency on bureaucratic quality for varying levels of democracy. Polity2 (democracy) scores are plotted on the x-axis, the marginal effect of a one-standard deviation change in each transparency measure is plotted on the y-axis. Point estimates are represented by the solid line, 95 percent highest posterior density intervals are represented by the dashed lines.

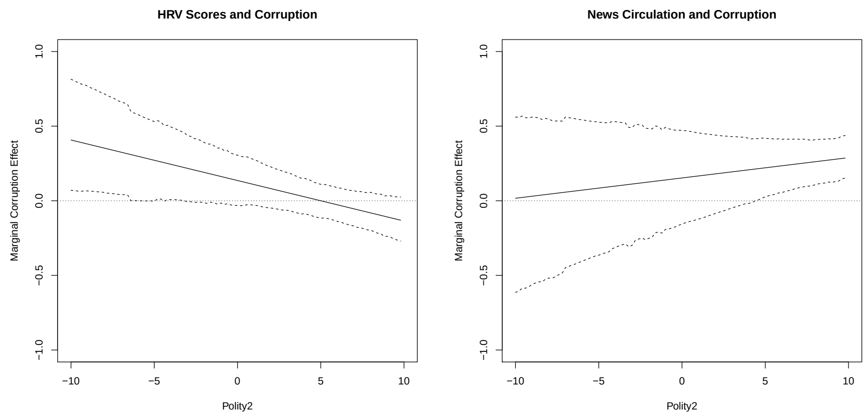
Figure 12: Marginal Effects of a Standard Deviation Change in Transparency on Law and Order



Plots of the marginal effect of transparency on law and order for varying levels of democracy. Polity2 (democracy) scores are plotted on the x-axis, the marginal effect of a one-standard deviation change in each transparency measure is plotted on the y-axis. Point estimates are represented by the solid line, 95 percent highest posterior density intervals are represented by the dashed lines.

mation, and are thus appropriate for use (without transformation) in statistical models. Because this index is based on the presence or absence of data from a commonly used source with broad coverage across time and countries, the HRV provides information on a far wider range of countries and dates than do

Figure 13: Marginal Effects of a Standard Deviation Change in Transparency on Corruption



Plots of the marginal effect of transparency on corruption for varying levels of democracy. Polity2 (democracy) scores are plotted on the x-axis, the marginal effect of a one-standard deviation change in each transparency measure is plotted on the y-axis. Point estimates are represented by the solid line, 95 percent highest posterior density intervals are represented by the dashed lines.

alternative transparency measures. Moreover, it captures a largely unmeasured form of transparency of great theoretical import.

In addition to constructing this index, we show that the HRV index is associated with governance outcomes even controlling for alternative forms of transparency. In particular, our index predicts the quality of governance in autocracies.

Of course, much remains to be done to more precisely document the mechanisms by which data dissemination affects government decision-making. Scholars should take greater care needs in specifying theoretical mechanisms, in constructing outcome measures, and in documenting the causal influence of transparency. Indeed, one direction of future work should reflect the decision of governments to release such information (for one example in this vein, see Hollyer, Rosendorff and Vreeland, 2011). This paper takes a necessary step in this direction by providing a valid measure of data dissemination.

References

Adserà, Alícia, Carles Boix and Mark Payne. 2003. “Are You Being Served? Political Accounatability and Quality of Government.” *The Journal of Law, Economics & Organization* 19(2):445–490.

Barro, Robert J. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14(1):19–42.

Berliner, Daniel. 2011. "The Political Origins of Transparency.".

Besley, Timothy. 2006. *Principled Agents? The Political Economy of Good Government*. Oxford University Press.

Besley, Timothy and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *The Quarterly Journal of Economics* 117(4):1415–1451.

Björkman, Martina and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *Quarterly Journal of Economics* 124(2):735–769.

Brambor, Thomas, William Roberts Clark and Matt Golder. 2005. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 13:1–20.

Broz, J. Lawrence. 2002. "Political System Transparency and Monetary Commitment Regimes." *International Organization* 56(4):861–887.

Brunetti, Aymo and Beatrice Weder. 2003. "A Free Press is Bad News for Corruption." *Journal of Public Economics* 87:1801–1824.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: The MIT Press.

Carmines, Edward G. and Richard A. Zeller. 1979. *Reliability and Validity Assessment*. SAGE Publication.

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143:67–101.

Chong, Alberto, Ana L De La O, Dean Karlan and Leonard Wantchekon. 2010. "Information Dissemination and Local Governments' Electoral Returns: Evidence from a Field Experiment in Mexico.".

Di Tella, Rafael and Ernesto Schargrodsky. 2003. "The Role of Wages and Auditing During a Crackdown on Corruption in the City of Buenos Aires." *Journal of Law, Economics & Organization* 269:269–292.

Duch, Raymond M. and Randolph T. Stevenson. 2008. *The Economic Vote: How Political and Economic Institutions Condition Election Results*. Cambridge University Press.

Fearon, James D. and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *The American Political Science Review* 97.1:75–90.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50:5–25.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publically Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2):703–745.

Fforde, Adam and Stefan De Vylder. 1996. *From Plan to Market: The Economic Transition in Vietnam*. Westview Press.

Fox, Jean-Paul and Cees A.W. Glas. 2003. "Bayesian Modeling of Measurement Error in Predictor Variables Using Item Response Theory." *Psychometrika* 68(2):169–191.

Gates, Scott, Havard Hegre, Mark P. Jones and Havard Strand. 2006. "Institutional Insconsistency and Political Instability: Polity Duration, 1800-2000." *American Journal of Political Science* 50(4):893–908.

Gehlbach, Scott and Edmund Malesky. forthcoming. "The Contribution of Veto Players to Economic Reform." *Journal of Politics* .

Gelman, Andrew and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Analytical Methods in Social Science Cambridge University Press.

Gelman, Andrew, Yuri Goegebeur, Francis Tuerlinckx and Iven Van Mechelen. 2000. "Diasgnostic Checks for Discrete Data Regression Models Using Posterior Predictive Simulations." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 49(2):247–268.

Gill, Jeff. 2008. "Is Partial-Dimension Convergence a Problem for IInference from MCMC Algorithms?" *Political Analysis* 16(2):153–178.

Herron, Michael C. 1999. "Postestimation Uncertainty in Limited Dependent Variable Models." *Political Analysis* 8(1):83–98.

Heston, Alan, Robert Summers and Bettina Aten. 2012. "Penn World Table Version 7.1." Center for International Comparisons of Production, Income and Prices at the Univeristy of Pennsylvania.

Hoffman, Barak D. 2011. "The Politics of Transparency in Kenya and Tanzania." Manuscript. Department of Government, Georgetown University.

Hollyer, James R., B. Peter Rosendorff and James R. Vreeland. 2011. "Democracy and Transparency." *Journal of Politics* 73(4):1–15.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2013*a*. "Data Disclosure, Accountability and the Facets of Transparency.".

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2013*b*. "Transparency, Protest and Political (In)Stability.".

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. N.d. "Replication data for: Measuring Transparency." http://dx.doi.org/10.7910/DVN/24274 UNF:5:Mj2tZ0rivXJPhuFqdQS0JA== IQSS Dataverse Network [Distributor] V1 [Version].

Islam, Roumeen. 2006. "Does More Transparency Go Along with Better Governance?" *Economics and Politics* 18(2):121–167.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Wiley.

Kiondo, Andrew. 1992. *Tanzania and the IMF: The Dynamics of Liberalization*. Westview Press chapter The Nature of Economic Reforms in Tanzania, pp. 21–42.

Knorr-Held, Leonhard. 1999. "Conditional Prior Proposals in Dynamic Models." *Scandinavian Journal of Statistics* 26(1):129–144.

Lebovic, James H. 2006. "Democracies and Transparency: Country Reports to the UN Register of Conventional Arms, 1992-2001." *Journal of Peace Research* 43(5):543–562.

Malesky, Edmund. 2008. "Strainght Ahead on Red: How Foreign Direct Investment Empowers Subnational Leaders." *Journal of Politics* 70(1):97–119.

Malesky, Edmund. 2009. "Gerrymandering Vietnam Style: Escaping the Partial Reform Equilibrium in a Non-Democratic Regime." *Journal of Politics* 71(1):132–159.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10:134–153.

Mmuya, Max and Amon Chaligha. 1992. *Towards Multipary Politics in Tanzania*. Dar es Salaam University Press.

Mmuya, Max and Amon Chaligha. 1994. *Political Parties and Democracy in Tanzania*. Dar es Salaam University Press.

Prunier, Gérard. 1996. "Somalia: Civil War, Intervention and Withdrawal." *Refugee Survey Quarterly* 15(1):35–85.

Reinikka, Ritva and Jakob Svensson. 2003. "The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture." Working Paper: Institute for International Economic Studies, Stockholm University.

Riedel, James and William S. Turley. 1999. "The Politics and Economics of Transition to an Open Market Economy in Viet Nam." OECD Development Centre Working Papers 152.

Rodrik, Dani. 1995. "Why is there International Lending?" NBER Working Paper.

Rosendorff, B. Peter and James Raymond Vreeland. 2006. "Democracy and Data Dissemination: The Effect of Political Regime on Transparency.".

Stone, Randall W. 2008. "The Scope of IMF Conditionality." *International Organization* 62(4):589–620.

Treisman, Daniel. 2004. "Stabilization Tactics in Latin America: Menem, Cardoso, and the Politics of Low Inflation." *Comparative Politics* 36(4):399–419.

Vreeland, James Raymond. 2003. *The IMF and Economic Development*. Cambridge University Press.

Williams, Andrew. 2009. "On the Release of Information by Governments: Causes and Consequences." *Journal of Development Economics* 89:124–138.