Transparency Squared: The Effects of Aid Transparency on Recipients' Corruption Levels

Zachary Christensen
Brigham Young University

Richard Nielsen
Harvard University

Daniel Nielson
Brigham Young University

Michael Tierney
College of William and Mary

Paper prepared for the 2011 meeting of the International Political Economy Society, Cambridge, Massachusetts, November 12-13.

Abstract

In the discussions surrounding aid effectiveness, donors have increasingly demanded lower levels of corruption among their recipients. Development advocacy NGOs and officials in recipient countries have responded by insisting on increased transparency from donors about aid projects. We ask the next logical question: Does donor transparency affect recipient corruption levels? If more information is available about donor activities at the level of individual projects in given recipient countries, less space ought to remain for bureaucrats or contractors to divert or misuse the committed funds. We measure aid transparency by quantifying the breadth of project information provided by donors. We explore the relationship between our measure of donor transparency and corruption levels in recipient countries. Drawing information from AidData, we use propensity score matching to account for the possibility that more transparent donors may be inclined to target less corrupt countries. The matching technique also enables greater confidence about the causal effects of donor transparency on recipient corruption. We find that more transparent aid in the aggregate appears to lead to lower corruption levels within recipient countries.

"Sunlight is said to be the best of disinfectants."[1]
– Justice Louis Brandeis

Introduction

Champions of aid call for more money to lift more people out of poverty (Sachs 2005; Barder 2006). Aid critics vigorously contest these demands, asserting that aid often – if not always – does more harm than good (Easterly 2006; Moyo 2009). Aid champions and critics agree about little, it seems. But on one point they sing in unison: aid should be more transparent (Easterly 2009; Barder 2009).

The consensus within the aid community holds that more and higher-quality information about aid should have many positive effects on outcomes valued by all interested parties. Transparency enables better coordination among donors. It makes greater specialization between aid agencies possible. It enhances the ability of recipient governments to effectively plan their budget expenditures. It allows stakeholders in both donor and recipient countries to engage the aid process more efficiently. It facilitates learning based on accumulation of research, shared evaluations, and best practices. And critically, greater information about aid permits citizens, the media, and watchdog groups to hold governments and their contractors accountable for how the aid is used.

It is this last point that motivates the present research. We ask a basic question that follows from the nearly universal enthusiasm for aid transparency: does donors' provision of more comprehensive aid information make recipient countries

[1] Louis D. Brandeis, 1932, "What Publicity Can Do," *Other People's Money,* chapter 5, p. 92. First published in *Harper's Weekly,* December 20, 1913.

less corrupt? The idea is relatively simple: donors' – such as the World Bank, the United States, and the European Community – making their aid more transparent should leave less space for recipient governments and their contractors to capture the aid through corruption or otherwise divert it for unstated purposes. As an aside, slightly more than one half of development assistance comes from multilateral organizations in the form of loans at rates below what the recipients can secure through international financial markets; the remainder of development assistance comes from bilateral aid organizations in the form of grants. Because some recipients are especially dependent upon foreign aid to supplement government revenue, and those same aid-dependent countries also tend to be the most corrupt on average, aid transparency may have a significant impact on dampening levels of corruption overall.

Of course, assessing this claim empirically may not prove as simple as stating the argument or telling illustrative stories. One problem is selection bias. That is, donors that value transparency may indeed make more information available about their aid. But at the same time transparent donors may also specifically target less corrupt countries as recipients of their aid. Such donors may thus extend the ethos of transparency to the selection of their beneficiaries. This would make it difficult to know whether transparent aid at time T1 caused a change in corruption levels at T2, or whether transparent donors just selected recipients based on current low levels of corruption or expected future low levels of corruption.

This potential problem requires careful attention to statistical modeling in a way that accounts for the possible confounding effect of recipient selection. We condition on this confounding factor using propensity score matching. Extracting information from nearly one million projects totaling $4.2 trillion in AidData from all major multilateral and bilateral donors, we first partition the sample into seven segments ranked according recipients' likelihood of receiving aid from the more transparent donors such as the European Commission, the Netherlands, and Sweden; we then perform regression analysis on each of the seven subclasses assessing the effects of aid transparency on corruption levels. We can then combine the results to produce an overall estimate of the effects of aid transparency on corruption.

Additionally, we check robustness through partitioning the data into subclasses reflecting the six different levels of corruption. We set a threshold for treatment indicating high transparency, and then we match "treatment" observations exactly to "control" counterparts using propensity scores generated by logistic regression. Discarding the observations not needed after the match, we perform regression analysis on the matched sample. The results produced by both methods strongly suggest that greater donor transparency leads to lower recipient corruption. In what follows, we situate this argument within the broader literatures on aid transparency and government corruption, develop the argument logically, describe our data and estimation procedures, perform the analysis, and explore the findings.

Literature

Prior research indicates that small improvements in government transparency have yielded significant positive outcomes. When local governments are held accountable by central governments, non-governmental organizations or citizens, the delivery of public policy improves to become more efficient. This has been found for development projects delivered by both non-governmental and governmental organizations (Olken 2007; Bjorkman and Svensson 2009; Duflo et al. 2008). Transparency, it seems, may decrease the incentive to provide inferior services.

Additional studies indicate that media scrutiny combined with government transparency especially decreases corruption. If the media is not able to report on the government in a way that the constituents can effectively interpret, the incentive structure for corruption remains largely unchanged. This work holds that transparency will only effectively improve aid if the proper groups are able to use the newly freed information to pinpoint problems and to support their demands for reform (Ferraz and Finan 2008; Francken et al. 2008; Eisenee and Stromberg 2007; Bailard 2009).

Alternatively, some analysts claim that transparency may have undesirable consequences. In the public sector, transparency may generate increased demand for particularly successful agencies or programs. These organizations have limited resources and must then divert energy from the substance of their work to try and accommodate new demands and requests for more information. In other situations, transparency may cause agents within development bureaucracies to seek projects that will make the organization look good rather than where need is greatest or where the agency has a comparative advantage (Prat, 2005; Gavazza and Lizzeri, 2007; Dranove, 2003). This body of work asserts that corruption should not necessarily be the first criterion for determining aid allocation. Need, economic growth potential, the requirements of a specific subset of the population, and other factors should determine aid allocation first; the corruption level of the recipient country ought to be secondary.

However, these objections assume that bureaucratic experts have a better sense of what they should do for the public good than the public itself. And as the literature on bureaucracies has amply demonstrated, it is probably safer to assume that bureaucrats are prone to pursue their own idiosyncratic preferences, which may include self-preservation, self-enrichment, a simplified task environment, or an ossified set of organizational principles (Moe 1990; Barnett and Finnemore 2004). While transparency may make the lives of government bureaucrats more difficult, the benefits of more public information seem to greatly outstrip the costs. Unfortunately, the empirical literature on the effects of aid transparency – particularly with regards to corruption – is small, so we are hopeful that the present study may help build scholarship in a promising, if underdeveloped, area of research.

Much better developed is the literature on the causes of corruption. Factors that appear to determine levels of corruption include electoral systems, the majority religion, GDP per capita, and colonial heritage. It also seems that cultural expectations and economic conditions can change the way that people respond to corruption (Treisman, 2007; Adsera et al., 2003; Persson, 2003; Serra, 2006; Billger and Goel, 2009; Swamy et al., 2001). Of particular interest to this study, Adsera et al. (2003) and Treisman (2007) find that countries with higher levels of newspaper readership have lower corruption scores. This supports the notion that information asymmetries facilitate corruption. The crucial difference between private and public information lessens when citizens can ask questions about government actions and journalists can report freely on potential instances of corruption without physical threat or fear of economic or legal persecution. It is this mechanism, among a range of other plausible alternatives, that likely causes corruption to decline in the face of increased transparency.

Argument

We thus hypothesize that more transparent aid from donors causes corruption to decline. As the share of aid that a given recipient collects from transparent donors increases, more is known about the details of specific development projects implemented within that recipient country. The greater degree of information enables interested parties – be they non-governmental watchdog groups, journalists, or ordinary citizens – to track the projects and verify that the money was spent for its intended purpose. When discrepancies between planned development activities and those actually implemented are identified, activists can trigger "fire alarms" that require politicians in the donor government – and perhaps even in the recipient government – to account for the project in question (see McCubbins and Schwartz 1984).

A micro-foundational logic underpins this argument. Governments and international organizations (IOs) make information about their actions public when it serves the interests of the politicians in office in the former case and their appointees to governing boards of IOs in the latter. Drawing on the work of pioneers in the agency literature (Coase 1937, Alchian and Demsetz 1972, Kiewiet and McCubbins 1990), we can imagine strong incentives for politicians and their IO appointees to hide their actions from voters, who form the ultimate "collective" principal with the authority to punish politicians and IO board members (Nielson and Tierney 2003, Lyne 2009). Hidden information enables politicians and IO delegates to pursue their own preferences – be they policy outcomes, longer terms, higher office, or more money – without meddling from their citizen employers.

Thus, the voters in the collective principal can achieve their objectives of better public policy with lower taxes only if they can learn where politicians and IO board members – and their subsequent agents, the bureaucrats – are slacking. Transparency enables voters to reduce the information asymmetry and achieve ends closer to their ideal points. Indeed, struggles over transparency mark the political histories of most developed countries and the recent experience of IOs, with the balance tilting toward voters as collective principal over time.

Aid transparency provides a particularly challenging test to this argument about voters' ability to reduce information asymmetries. Indeed, voters in donor countries do not receive the benefits of foreign aid directly and thus cannot monitor government policy in the same way, for example, they might notice the quality of their nearby roads, schools, or hospitals. And IOs are yet one more step (or more) removed from voters, who are spread across numerous member countries. In both cases monitoring of foreign aid can only happen at great distances, and the primary beneficiaries cannot directly influence the incentives of their benefactors. And this likely reduces the interest and effectiveness of voters in monitoring foreign aid outcomes (Martens et al. 2008).

Yet there remains a great deal of variance in the degree of transparency among different aid portfolios at both multilateral and bilateral donors, suggesting that incentives for aid transparency do exist in some countries and for some multilateral aid organizations. And it is precisely the effects of this variance that most interest us here. When donors are transparent about their aid – even if this transparency is merely an artifact of standard bureaucratic procedures to make organizational activities matters of public record – the published information enables citizens, the media, and watchdog groups to scrutinize donor portfolios in a way that they cannot when aid hides in the shadows of vague bureaucratic reporting.

We argue that knowing that their aid is being scrutinized – or at least watched with more care – likely motivates politicians and, by extension, IO boards and aid bureaucrats at transparent donors to monitor the process and outputs of their own aid projects more closely. In such settings we expect, for example, the development of procedures for procurement and contracting that make the siphoning of aid to anonymous shell corporations more difficult. Such donors also ought to take care to target their aid more specifically and perhaps even earmark some of it for improving recipient-country governance, including projects that combat corruption. And we ought to see occasions where watchdog groups or journalists use public information about the aid to shame government officials or IOs and demand reforms (see Nielson and Tierney 2003).

When thinking about this, it may prove useful to imagine the mindset of a politician or IO board member whose job it is – given a ministry position, legislative committee assignment, or international appointment – to allocate foreign aid. The official may operate under a strong transparency regime, where domestic law or IO bylaws dictate that thorough reports about aid projects are published in a timely fashion. Or not. And it is this variance that proves key. If the aid portfolio is transparent, with projects thoroughly described and categorized according to international standards, the politician can anticipate that others in the aid community, including NGOs or newspaper reporters, can track the aid, identify problems, and raise concerns publicly.

Indeed, scandals involving foreign aid gone awry sometimes become quite prominent, occasionally even prompting serious reforms at the aid organization in question (see Nielson and Tierney 2003). When indigenous people dressed in traditional garb show up to hearings decrying the effects of aid projects on their communities, they grab headlines. When these indigenous people are joined by environmental NGOs (representing voters) who care about environmental protection and taxpayer organizations (representing voters) concerned about their money being wasted, politicians and their appointees start to pay attention. Aid officials, aware of this potential, should take greater care with the aid portfolio to the degree that the public can learn about aid projects and punish them for wasting money, destroying the environment, or doing other unpopular things in their name. When aid is vaguely described or broadly categorized, politicians and IO board members face reduced incentives to monitor the projects to ensure that their intended purposes are realized. After all, who will ever know?

In this paper we do not attempt to explain why some donors are more transparent than others. We merely note that the variance in donor information generates different incentives for politicians and bureaucrats working under the alternative transparency regimes. More transparent regimes generate greater incentives for politicians and their agents to undertake better scrutiny of projects funded by their organizations. Corruption should go down in recipient countries that receive significant aid from these transparent donors. On the other hand, less transparency enables greater hidden action. Corruption should logically follow.

Data and Measurement

To undertake an empirical analysis of the effects of donor transparency on recipient corruption, we construct a time-series cross-national dataset covering 95 countries between 1989 and 2004 for a total of roughly 1300 country-years. This timespan was selected because of current data limitations on key explanatory variables.

Dependent Variable: Recipient Corruption Levels

We use a measure of corruption developed by the International Country Risk Guide (ICRG), which ranks countries on perceptions of corruption using a 6-point scale (with 1 indicating severe corruption and 6 indicating no corruption). The ICRG, like most other measures of corruption, surveys experts each year to obtain scores on perceptions of corruption in all of the countries included. To capture broad trends in corruption rather than year-to-year fluctuations, we use a three-year rolling average of ICRG corruption scores, taking care that all of these years occur in time after our independent variables to avoid the possibility of direct reverse causation.[2]

This measure of corruption is attractive because of its relatively broad coverage among developing countries that are recipients of aid as well as its substantial temporal coverage (from 1984 onward). However, some countries are missing from this dataset, presumably for non-random reasons. In particular, we might expect that the ICRG omits countries that hold little or no interest to western businesses and investors, and the omissions may be correlated with corruption. In future research, we plan to use multiple imputation to estimate corruption scores for developing countries that are not in the ICRG dataset, but for now we resort to list-wise deletion of missing country-years.

[2] We also test moving averages of 1, 2, 4, and 5 years and obtain similar results to those presented below.

Independent Variable: Donor Transparency

Our goal is to measure donor transparency regarding aid projects in a way that captures our intuition that more information about aid reduces possibilities for corruption of the aid process. At the project level, our goal is to measure the amount of information available about the size, scope, target, and objectives of an aid project. Conveniently, the AidData 1.91 database provides information for assessing donor transparency at the project level – because AidData represents an attempt to collect comprehensive project-level data on the universe of aid, we can use the availability of information in AidData as a proxy for the transparency of a given aid project.

We underscore that this is merely a proxy. A full assessment of the transparency of given donors would do more than simply tally the number of fields reported for each project. One could independently assess the amount of information on a donor website, in donor annual reports, in project documents made public, in agency evaluations, in the percent of public project audits, in the length of project descriptions in publicly available datasets, and so on.

That said, we argue that our simple measure is a good starting point. Throughout its history, the Development Assistance Committee (DAC) of the Organization of Economic Cooperation and Development (OECD) has required member countries to complete an annual survey and submit responses about their aid activities. Most – though by no means all – governments comply at least at the minimum required level. But other governments provide more extensive information than required.

When we have inquired why some donors provide greater information than the DAC requires, the response is that such donor organizations desire to be forthcoming about their aid *and* the information is readily available in their aid information systems. The ready availability of such information generally results from the reporting requirements at the country level. Politicians and, ultimately, voters monitoring the more transparent organizations apparently have succeeded in extracting more information about the aid. Greater breadth of aid information thus generally results from more stringent reporting requirements placed by the principals on the aid organizations, be they bilateral or multilateral.

Thus, to construct a project-level measure of transparency we identify 55 fields in the AidData 1.91 dataset that we surmise ought to contribute to the transparency of a project. This information was either reported by donors to the OECD's Creditor Reporting Service (the case for most bilateral donors) or AidData staff members collected the information from donor documents or websites (the case for most multilaterals). While the number of fields reported in AidData likely does not capture all available information about the aid, it does likely encompass most of the information that is readily accessible and thus effectively in the public domain. In the count of 55, we include fields that identify projects, project managers, project purposes, financing details, and implementing agencies. The specific fields used are listed in the appendix. For each project in the database, we count the number of these fields that are populated and use this as our project-level measure of transparency.

In Figure 1 we chart the median number of fields reported for a selected set of donors for their aid projects allocated between 2003 and 2008. There is significant variance in the number of fields reported, with the ordinal ranking conforming generally – though not perfectly – with other metrics of aid transparency (see Easterly and Pfutze 2008, Birdsall and Kharas 2010). Figure 2 demonstrates for a smaller select set of donors that the median number of fields reported has changed over time – even in recent history since 2003 – for several donors.

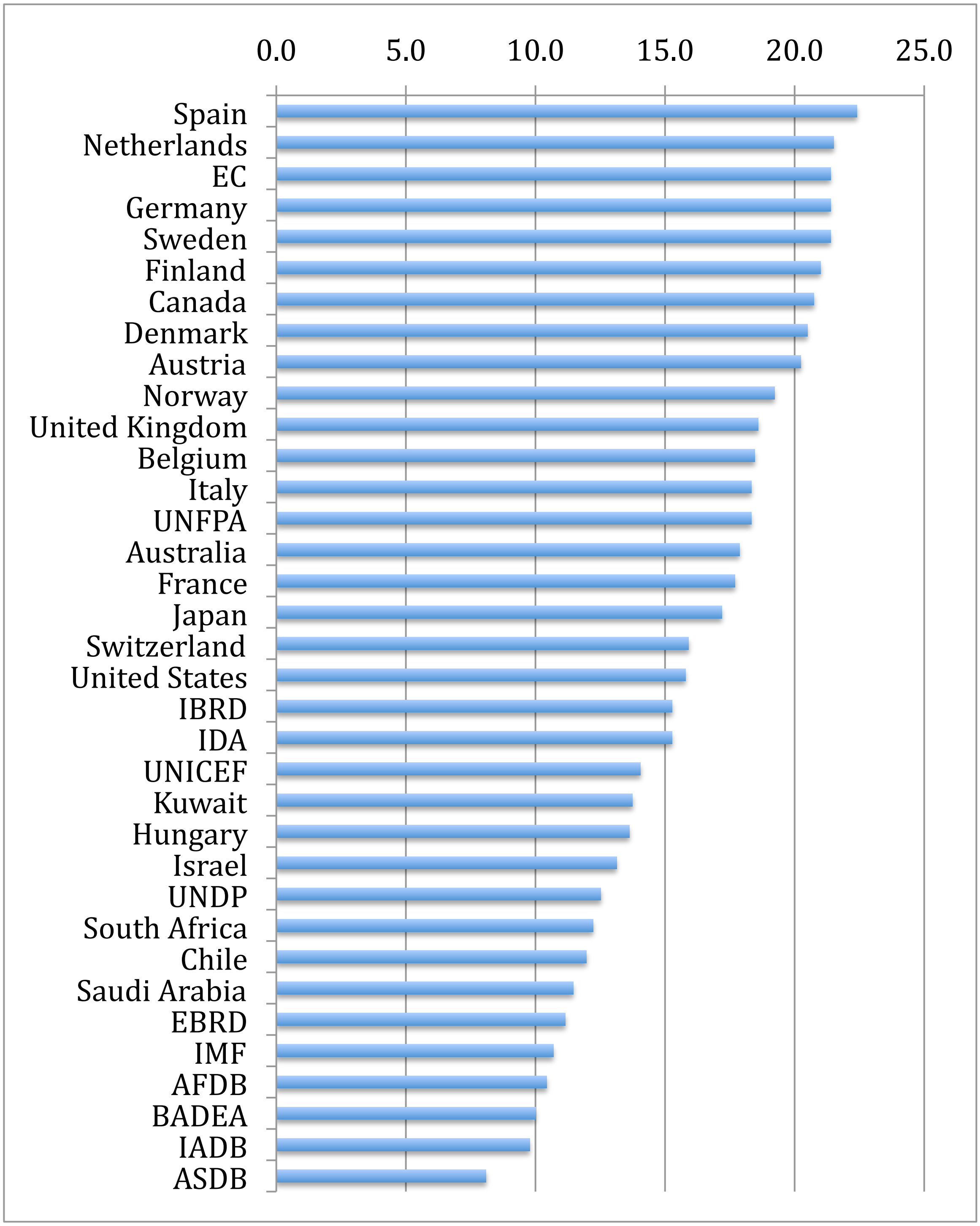

Figure 1: Ranking of selected multilateral and bilateral aid organizations by median number of fields listed in AidData, 2003-2008.

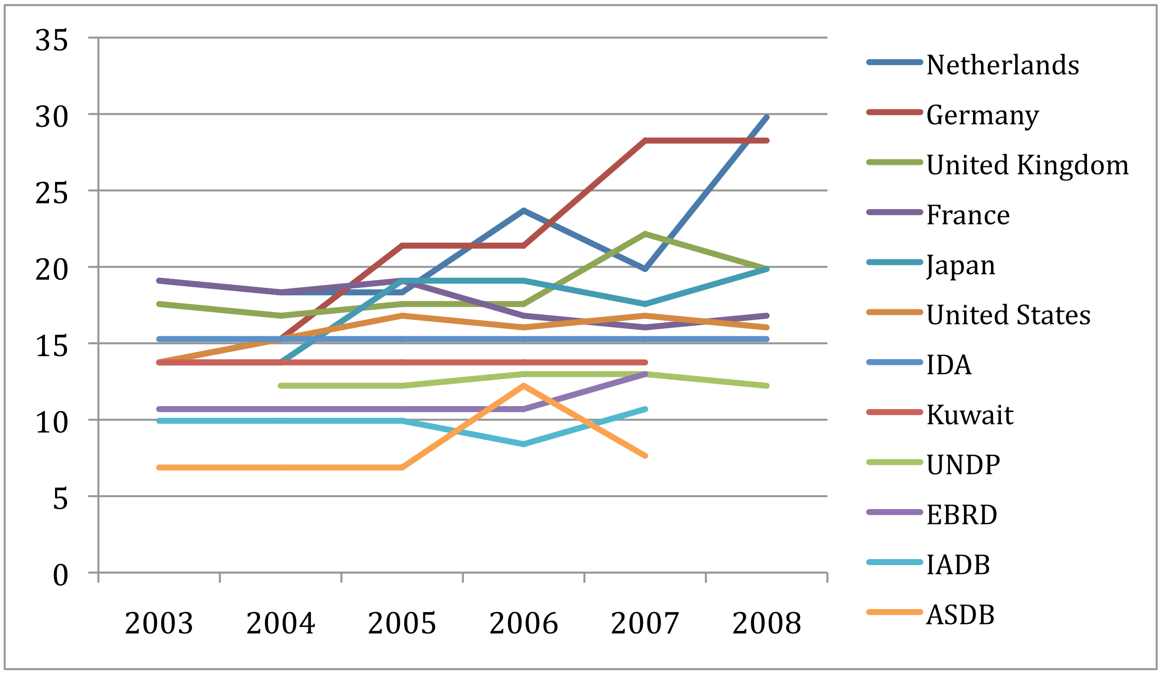

Figure 2: Median number of reported AidData fields for selected donors from 2003 to 2008.

Skeptics may reasonably worry that some of these fields are more important than others. For example, our intuition suggests that long descriptions of projects and information about the channel through which project are delivered might be especially important bits of information for any third party attempting to hold governments or contractors to account. Thus, they are, arguably, more important indicators of transparency than some of the other fields. However, weighting fields in this way is likely to be contentious and there is little *a priori* information we can draw upon to determine how the fields should be weighted. We hope to administer a survey of aid practitioners asking them to rank each of these fields in importance, which would allow us to place more weight on important fields. For the moment, our measurement implicitly presumes equal weights – a contestable assumption that each of the fields is equally important.

From our project-level definition of transparency as the number of populated fields, we develop a measure of aid transparency for each recipient-year. To do this, we simply examine all of the projects received by a recipient country in a given year and tally the median number of populated data fields for all of these projects (we use the median, rather than the mean, to make the measure robust to outliers). This variable, *Median Coverage*, is our primary variable for measuring aid transparency at the recipient-year level. We also believe that multi-year trends in aid transparency are more important for determining corruption outcomes than year-to-year fluctuations, so we calculate the 5-year moving average of *Median Coverage*, which we call *Median Coverage (5 year)*.[3]

This measure may be counterintuitive because aid transparency is typically conceptualized in terms of donors: "How transparent is Danish aid relative to aid from the United Kingdom?" However, for our purposes, we need a measure that captures a slightly different question: "How transparent is the aid received by Zimbabwe relative to the aid received by Uruguay?" Our measure answers this recipient-level question.[4]

[3] Our choice to average over five years is based on a loose theory that aid transparency needs some time to have any appreciable effect on recipient corruption. We test alternative moving averages of 2,3,4, and 6 years and find very similar results to those reported below.

[4] If we had started off by measuring donor transparency and then tried to map this to recipients, we would be relying on the assumption that donors are equally transparent with different recipients, which may be a faulty assumption.

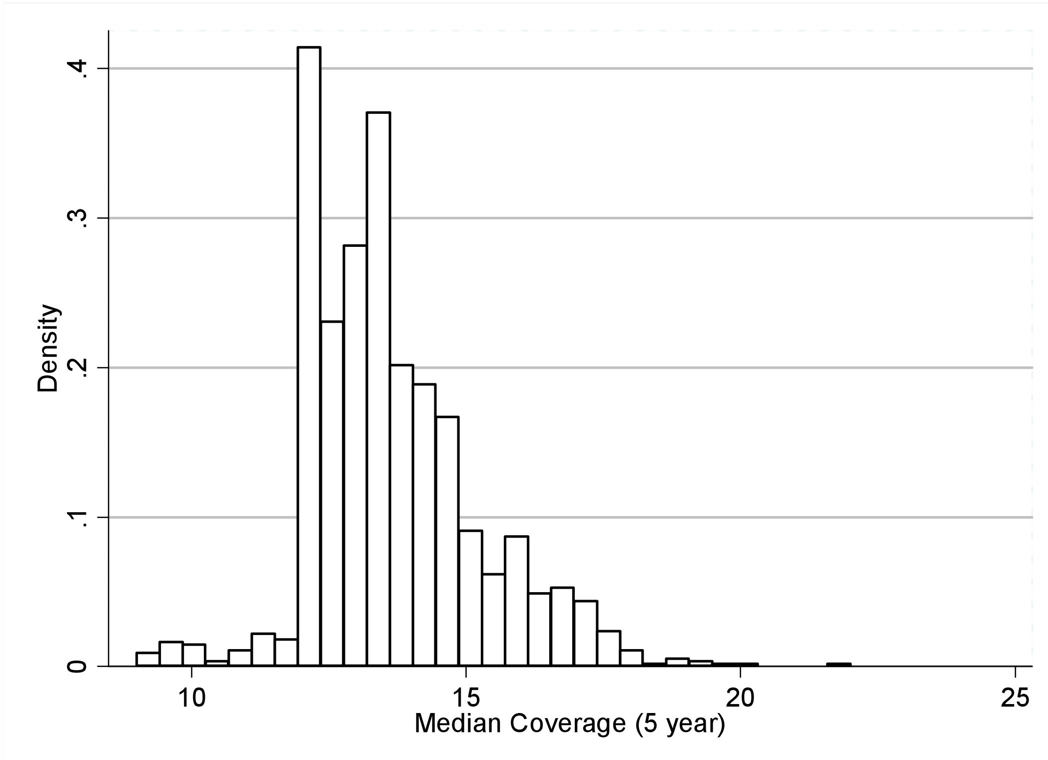

Figure 2: Distribution of *Median Coverage (5 year)* for the 1313 recipient-year observations used in our sample.

In practice, relatively few of the 55 fields we consider have coverage. In the sample for which we estimate models, the minimum observed value of *Median Coverage (5 year)* is 9 and the maximum is 22, with a mean of 13.6. The full distribution of *Median Coverage (5 year)* is shown in Figure 1. Descriptively, the recipient country-years with the worst data coverage are Saudi Arabia, Venezuela, and Taiwan in the early 1990s. If we average median coverage over all years in the sample, then Kuwait and Saudi Arabia have received the least transparent aid in our dataset while South Africa, Albania, and Serbia have obtained the most transparent aid. If instead we look at average transparency by year, there is a clear trend of increasing transparency over time (within the time-frame of our study), with the lowest average data coverage in 1991 and the highest in 2003.

Matching Design

Our prior expectations about aid transparency and recipient corruption lead us to the conclusion that confounding, also called endogeneity, may compromise our empirical analysis if left uncorrected. Specifically, it is implausible that aid transparency is randomly assigned to recipients; instead it seems most likely that aid transparency is conditioned on the characteristics of recipients and their donors. Most problematically, we might expect that donors consider recipient corruption when they decide how transparently to report information about an aid project. It may be the case that transparent donors prefer to give their aid to less corrupt countries, potentially inflating a naïve estimate of the effect of aid transparency on corruption. However, it could also be the case that donors fear capture of their aid in corrupt environments so they are *more* transparent when they give aid to corrupt countries in order to limit the ability of the recipient to simply funnel aid into elite bank accounts. Or there may be some more complicated relationship between recipient characteristics and subsequent aid transparency.

We use propensity score matching methods, adapted for our time-series cross-sectional data structure, to account for these possible confounds. Propensity score matching is a nonparametric method for processing data in preparation for subsequent estimation of causal effects. In effect, matching tries to fix the "broken" experiment that gave rise to the data we observe. If aid transparency had been assigned randomly, estimation of the causal effect of aid transparency on subsequent recipient corruption would be relatively trivial. Our objective with matching is to account for the systematic differences between countries that receive different levels of aid transparency by matching countries that have similar values of the factors that determine aid transparency. We then assume that after accounting for the systematic parts of the aid allocation process, there is still some randomness in the process such that identical countries might still get somewhat different levels of aid transparency. We exploit this randomness to estimate the effect of increased aid transparency on corruption *for countries that were similar prior to receiving the aid*.

There are numerous matching methods and vigorous debate rages about which are best for estimating causal effects. We present the results of two different styles of matching: (1) *subclassification on propensity scores* and (2) *exact matching on past corruption followed by one-to-one propensity score matching*. With both, we find that aid transparency has a positive effect on subsequent recipient corruption.

Subclassification on Propensity Scores

We perform our first matching procedure and subsequent estimation in the following steps, each of which we discuss in detail below.

- Define the treatment, define the units of observation, and define the target causal effect.
- Estimate a model for how treatment is assigned to recipient-years.
- Use this model to calculate propensity scores for each unit.
- Divide the sample into subclasses based on the propensity scores.

- Estimate a regression model within each subclass to estimate the treatment effect.
- Combine the subclass estimates into a single estimate – the average linear effect across all units.

We define the unit of observation to the country-year and our treatment is the variable *Median Coverage (5 year)* defined above. The outcome of interest is subsequent corruption scores, averaged over the three subsequent years. Thus, the causal question motivating our analysis is "what is the effect of the past five years of aid transparency on the next three years' level of corruption?" [5]

To account for the non-random allocation of aid transparency, we estimate a statistical model predicting the level of transparency of the aid that a given country receives in a given year. As we measure it, *Median Coverage (5 year)* is continuous so we use a hierarchical linear model with recipient-level random intercepts.

The choice of predictors to include in this model of recipient-level aid transparency is very important. In particular, we need to include every variable that influences a recipient's level of aid transparency that might also be a correlated with subsequent corruption. To elaborate this set of covariates, it is helpful to think about the possible sources of confounding.

First, it could be that aid transparency is assigned based on a recipient's previous levels of corruption and that previous corruption is correlated with

[5] Recall that our Median Coverage (5 year) is averaged as well. For example, we are estimating the effect of aid transparency between 1990-1994 on corruption levels in 1995-1997.

subsequent corruption. We therefore condition on the 5-year lag of corruption in the propensity score model. (Note that all variables in this model will be lagged 5 years so that they are all pre-treatment.)

Less directly, we might worry that international norms about corruption and transparency influence both donors and recipients. In particular, if international actors are shaming donors for being less than transparent in their aid giving and also shaming recipients about having corrupt business practices, then this might induce spurious correlation between transparency and subsequent recipient corruption. We account for this by including the average level of *Median Coverage (5 year)* across all recipients as a covariate in the propensity score model. This accounts for any time trending in average transparency.

The final major path for endogeneity to operate is through the donors themselves. First, different donors that are more or less transparent in general might select to give aid to different types of recipient countries for reasons such as colonial ties, trade ties, alliances, and so on. Many of these reasons might also be correlated with recipient corruption outcomes; for example, colonial ties that influence who a recipient's primary donors are might also have shaped its legal origins and legacy and might thus be correlated with current levels of corruption. In this case, it is helpful to think about the experiment we might have wished to run: we would have liked to have the same donors give aid to the same types of countries but randomize the amount of information reported about the aid, making it either high or low. In the absence of randomization, our goal is to make comparisons between recipients that have similar donors and similar characteristics, but that for (arguably) random reasons had differing levels of median aid transparency.

We thus include as covariates the amount of aid from each of the 17 bilateral OECD donors. The intuition behind this is that if we are comparing recipients that have similar donor profiles then subsequent differences in corruption levels between these recipients are not attributable to donors selecting to aid certain recipients over others. Although conditioning on the donors could solve the problem entirely, we also condition on the other variables along these potentially endogenous causal pathways. We include indicator variables for former colonies of the UK, France, Portugal, Spain, Belgium, the Netherlands, and former Soviet satellites. We also condition directly on recipient trade flows with most of the OECD donors, as well as an indicator for states that have a military alliance with at least one donor country.

Finally, we round out the set of covariates with a standard collection of variables from the corruption and aid allocation literatures measured at the recipient level: democracy, respect for political rights, GDP per capita, population, regional indicators, and an indicator for the Cold War. We acknowledge that this may seem like a "garbage-can" regression. This is in fact the goal – by conditioning on as many pretreatment covariates as possible, we decrease the likelihood that our results are spurious.

Propensity Score Results and Subclassification

The results of the propensity score model are shown in Table A1 of the Appendix. The coefficients on the predictors should probably not be interpreted causally (note that a propensity score model does not need to be causal), but the estimated effects are suggestive. We find that aid transparency is clearly not randomly assigned – there are substantial and meaningful differences between country-years that receive different levels of aid transparency.

First, we note that all three of the categories of potential confounders we described above seem to matter for predicting levels of aid transparency. Past corruption is a strong *positive* predictor of subsequent aid transparency. This is not a causal model so we should be careful with interpretations beyond correlation, but this does suggest that donors may take care to spell out more details about their projects that go to corrupt countries. This is promising for our results – if past corruption is the main confounder, then if anything, a naïve model will *underestimate* the effect of aid transparency on corruption. Second, we note that aid transparency is clearly time-trended and that the average global level of transparency strongly predicts transparency at the recipient level. By controlling for this variable, we avoid the possibility that this trending biases our results by also being correlated with trends in corruption levels.

Finally, we note that some aid donors appear to be less transparent than others, so recipients that receive more aid from these donors have lower *Median Coverage,* although almost none of these differences are statistically significant. Other types of ties to particular donors are significant – in particular, strong trade ties to a few of the smaller OECD donors decreases aid transparency. We also find that alliances with donors decrease transparency and, more surprisingly, that more democratic countries get less transparent aid (although we are controlling for corruption, so this is a partial effect).

Having estimated a model predicting aid transparency at the recipient-year level, we now use it to generate propensity scores for each recipient-year. In the case of a binary treatment (ex, "transparent or not"), propensity scores are typically calculated by estimating a logistic regression predicting the treatment status of each unit. The predicted values from this logistic regression (ranging from zero to one) are then used as the propensity scores for each unit. In this setting with a simple binary treatment, the propensity score can be intuitively interpreted as the probability that a unit is assigned treatment. These propensity scores can then be used to match or subclassify the units in the sample; by comparing units that have similar propensity scores, analysts will be comparing units that had similar probabilities of receiving treatment, so actual assignment to treatment or control will be (hopefully) up to chance.

Propensity scores are also a linear combination of the covariates in the propensity score model, so propensity scores provide a one-dimensional summary of the many-dimensional differences between units. Thus, within a matched sample or within matched subclasses, we are comparing units that have relatively similar values of the covariates predicting treatment. In short, rather than making extreme comparisons, the use of propensity scores allows us to compare "like to like."

Our treatment is a continuous variable so we cannot use standard matching methods developed for binary treatments without arbitrarily dichotomizing a continuous variable. Instead, we use a relatively new set of methods which use generalized propensity scores to create balance with multi-valued or continuous treatments (Dyk and Imai 2004).

The basic intuition is the same as for propensity scores extracted from a logistic regression predicting binary treatment. However, we have used a linear model to estimate the model predicting aid transparency. As with the logit case, we simply take the predicted values for each unit as the generalized propensity score or "balancing" score. This no longer can be interpreted as the probability of treatment; instead it is the predicted level of aid transparency for a given country-year. Following the logic of propensity scores in the logit case, country-years that have similar balancing scores will have similar values of the background covariates. The major difference between propensity score methods with binary treatments and the generalized propensity score methods we use for our continuous treatment is that we require the assumption of linearity of the causal effect.

We collect the predicted values from the linear regression and subclassify the dataset into seven equally sized subclasses of observations based on these propensity scores. This is roughly in keeping with the rule of thumb developed by

Cochran (1953) and extended by Rubin (1973) that using 5 or more subclasses removes roughly 90 percent of the bias for many applications.

Results

Within each subclass, we estimate the subclass causal effect by estimating a linear regression in which the outcome is *Corruption* and the treatment is our *Median Coverage (5 years)*. We include some of the same covariates from the propensity score model: the global average of *Median Coverage*, Corruption lagged 5 years, Democracy, Political Rights, Aid levels, Population, GDP per capita, and regional dummies. We also include country-level random intercepts. Because we are estimating this model on subclasses of the data specified by propensity score ranges, the results within each subclass will be less model dependent than the same regression estimated using the entire sample (Ho et al. 2007). If we have successfully accounted for the factors that influence aid allocation,[6] then the estimated coefficient on *Median Coverage* can be interpreted causally, as the effect of aid transparency on recipient corruption within the subgroup. We can then calculate the overall linear average effect by calculating the weighted average of the subclass estimates, with the weighting corresponding to the number of observations in each subclass.

Figure 2 shows the results of the seven models, one for each subclass, as well as the combined causal estimate (the average treatment effect) for all units, and the

[6] We cannot rule out that some unobserved and un-theorized variable is possibly driving our results, although we consider it unlikely.

naïve estimate that we would have obtained if we used the hierarchical model *without* subclassification on propensity scores. The coefficients shown are only the coefficients on *Median Coverage*. We do not report the coefficients of the control variables in these models because, after matching, these other coefficients have no causal interpretation.

The results show that aid transparency, as measured by data coverage, has a positive and statistically significant effect on subsequent donor corruption. In substantive terms, the effect is surprisingly large – if donors were to move from reporting 10 fields to 20 fields, the average recipient would experience an increase in their corruption score of roughly 0.3, which is non-trivial on a 6 point scale (remember that higher numbers mean less corruption on this scale). If the effect were much larger, we would begin to question the credibility of our analysis. Substantively, this result suggests that the general upward creep in donor transparency has probably had a small but detectible effect of reducing corruption in countries that receive relatively transparent aid.

It bears noting that the effects of donor transparency on corruption levels do not appear to be uniform across the subclasses. That is, for the recipients in the highest two subclasses most likely to receive transparent aid, more transparent aid within those subclasses appears to be associated with greater corruption. These results should be interpreted more cautiously than the overall finding that donor transparency leads to less corruption. Making inferences *within* the subclasses is more model-dependent because the strength of the matching lies in defining the subclass boundaries. Still, the trend toward smaller effects in the higher subclasses is provocative. It may be the case that those recipients with the most transparent aid have little practical room to improve: they are already at the lower levels of corruption and can only worsen from their current position. A slightly different interpretation might be that transparency has the biggest effects in the aid recipients where it is most surprising – where "surprising" means that a state had a low "propensity" to get transparent aid but received more transparent aid than was expected. More analysis will be necessary to tease out the implications of the heterogeneity across the subclasses.

We note also that the positive average effect we find might be somewhat surprising: can donors really shift aggregate measures of recipient corruption by making more information available about aid projects? One possibility, referenced above, is that this information gives watchdog actors a substantial leg up when trying to police governments to deliver promised aid in its intended form. While this information may represent just a few cells in our data matrix, it may be that even one additional cell gives aid constituents and activists additional leverage to fight the corruption of aid projects. And especially in countries that are highly aid dependent, reducing aid corruption may substantially reduce aggregate corruption.

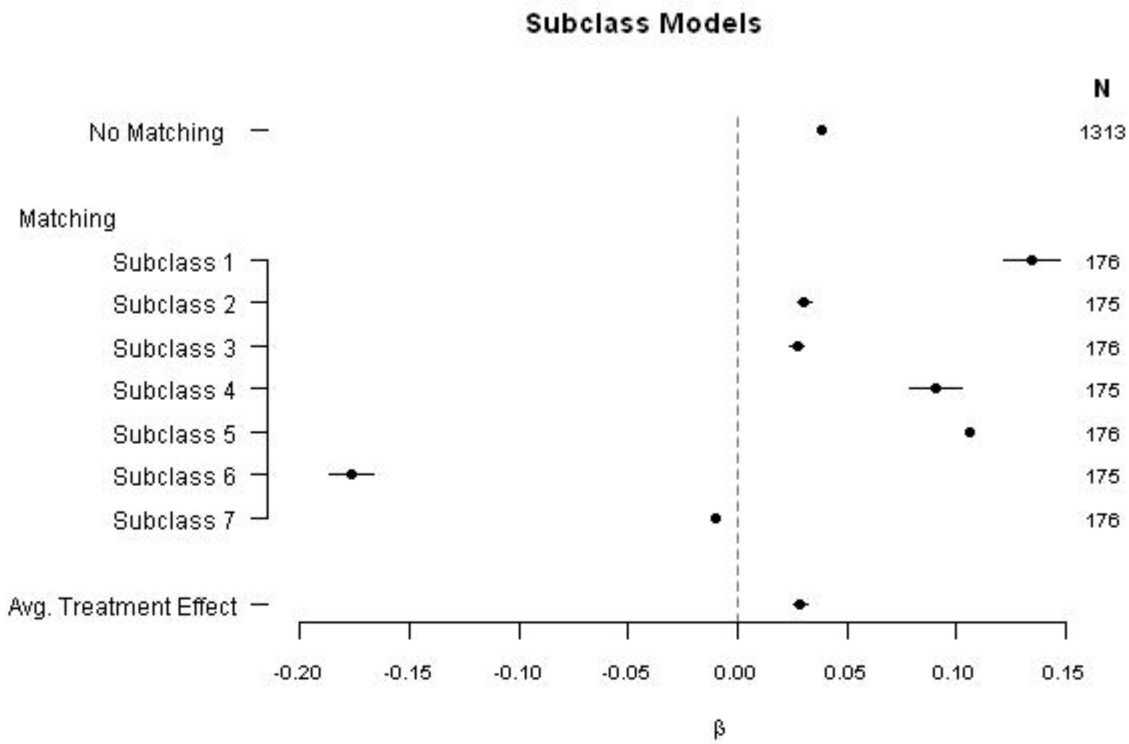

Figure 2: Estimates of the average effect of the past five years of aid transparency on corruption levels in the subsequent three years. The "No Matching" is obtained by estimating the hierarchical model described in the text using the entire sample without any matching or subclassification. Low numbered subclasses had the lowest predicted levels of aid transparency while high numbered subclasses had high predicted aid transparency. The estimated causal effect within each subclass is shown and the average treatment effect (ATE) is the average effect for the seven subclasses, weighted by the number of observations in each subclass. "N" shows the number of country-years in each subclass.

Another possibility is that aid transparency, as measured by data coverage, is really picking up on an underlying dynamic between donors and recipients. It may be that "donor transparency" is best thought of as a latent characteristic that can be measured with changes in data coverage but that operates primarily through other channels. For example, it may be that "anti-corruption-minded" donors report more information about their projects, but that they reduce recipient corruption via more direct means of pressure – coercion, the threat of withholding aid, anti-corruption training, etc. If this story is correct, then data coverage is a proxy for donor transparency, not a direct measure of it, and this substantial effect is not just the effect of additional data cells. Our data are unable to adjudicate between these two possibilities, though for theoretical reasons we prefer to interpret the results for the effects of *Median Coverage (5 year)* as a proxy: donor transparency for given projects merely reflects the greater level of scrutiny the aid organization faces overall, which as a whole induces more careful attention to where and for what purposes the aid money flows.

These results are all conditional on our somewhat arbitrary choice to use the past 5 years of *Median Coverage* to predict the next 3 years of recipient corruption. However, our results are not sensitive to this choice. In particular, we tested models that used different numbers of past years of *Coverage* (2, 3, 4, 5, and 6 years), and different numbers of years of corruption data (1, 2, 3, 4 and 5 years). We re-ran the entire matching and estimation procedure for each of the 25 possible combinations. The results are all similar to those presented here, with a number of models estimating even larger effects. The smallest estimated effects are roughly 0.02 (meaning that each additional field increases the average corruption score by 0.02 points on a 6-point scale) and the largest effects are roughly 0.065. We also find trends in the estimated effects that suggest they are consistent with out our causal logic. The estimated effect of transparency gets larger as we use more years to construct the average level of transparency. This suggests that prolonged exposure to transparency has larger average effects than short exposure, as we might expect. We also see a smaller trend towards larger effects in models that include more years of recipient corruption data. Together, these results imply that the effect of transparency on corruption deepens over time.

Exact Matching on Past Corruption Followed by One-to-one Propensity Score Matching

Our analysis is open to several criticisms. First, our approach to matching is hardly a common one in the applied matching literature. We have chosen to take on the additional assumptions necessary to use matching with a continuous treatment; some applied researchers might instead make the choice to dichotomize the continuous treatment and use a matching method other than subclassification. We are concerned about dichotomization because arbitrary dichotomization of a continuous treatment *ensures* violations of the key SUTVA assumption that there are no hidden versions of treatment.[7] Once we dichotomize aid transparency at some arbitrary threshold, a "1" may mean an aid transparency level of 14 for one country-year and of 20 for another country year.

[7] SUTVA is the Stable Unit Treatment Value Assumption that is necessary for most modern styles of causal inference. Formally, this assumption requires that the potential outcomes of treated and control units be conditionally independent given the covariates. Informally, this implies that there are no interactions between units and that there are no "hidden" versions of treatment.

However, we acknowledge that subclassification has weaknesses. In particular, although it bounds the imbalance on the covariates within each subclass, it relies heavily on the linear model to remove any remaining selection effects operating *within* subclasses. Intuitively, observations within a subclass are close, but not as close as if we had performed one-to-one matching.

To examine the sensitivity of our findings to alternative matching methods, we use the following procedure. First, we dichotomize the treatment variable, coding everything with *Median Coverage (5 years)* greater than 15.5 as "1" and everything with fewer fields as "0." We chose 15.5 arbitrarily (and without peeking!) because it is roughly the 75th percentile of *Median Coverage (5 years)*. This new dichotomous treatment might be informally conceptualized as an indicator variable for "exceptionally good coverage." We then perform a two-stage matching procedure. First, we are again concerned about past corruption as the primary confounder in our analysis, so we match *exactly* on the 5-year lag of corruption scores. This means that when we eventually fit a model to the data, we will be comparing countries that had *exactly* the same distribution of previous corruption prior to getting different levels of aid transparency/data coverage.

Then, within the 6 strata defined by the exact matching (one for each value taken on by the corruption score), we estimate a propensity score model identical to the one used for subclassification except that this time we use logistic regression because the new treatment is binary. Rather than subclassifying, we instead match treated units to the control unit that has the most similar propensity score, which gives us one-to-one matches for the treated units. There are fewer treated units than controls (by design) so we discard the control units that don't get matched to the treated units (meaning that they look substantially different from the treated units).[8] We then recombine the matched data from the 6 subclasses into a single, more balanced dataset and estimate the same regression model as before on the matched data using the binary treatment indicator.

Figure 3 shows the results from using this matching procedure. We again find that increased coverage has a positive effect on subsequent corruption levels. Substantively, the effect is again relatively large: receiving "treatment" of high *Median Coverage* leads an improvement in corruption score of roughly 0.25. We note that this result is on the same order of magnitude as the finding in the previous section, suggesting that different matching methods recover the same result.

This matching method has the weakness that we dichotomized at an arbitrary point. And, indeed, we find different results at different cut-points. In particular, we estimated models where we dichotomized at the 65th, 70th, 75th, 80th, 85th, and 90th percentiles of *Median Coverage.* The point estimates remain positive, but for some cut-offs (the 65th, 80th and 85th percentiles), the 95 percent confidence intervals around the estimated effect of dichotomous aid shocks contains zero. This is in part due to the decreases in sample size that occur from iteratively narrowing the definition of "treatment". Moreover, we are concerned about some of the clear

[8] Some people are uncomfortable with the way that typical matching methods discard data. More data is only better for causal inference if the data is comparable on relevant covariates. If matching discards observations, it is because these observations were irrelevant for estimating causal effects.

violations of the core assumptions of causal inference that are necessary for dichotomizing the treatment in this way – in particular, the point estimates for “treatments” defined by different cut-offs are actually estimating *different* target causal effects, even if they are unbiased. Because of these potential problems with the more standard matching analysis, we take some comfort in the fact that both matching methods seem to agree on the direction of the relationship. The weight of the evidence points to the encouraging conclusion that increased transparency in aid data leads to less corruption in aid projects, but we hardly claim that our findings are definitive.

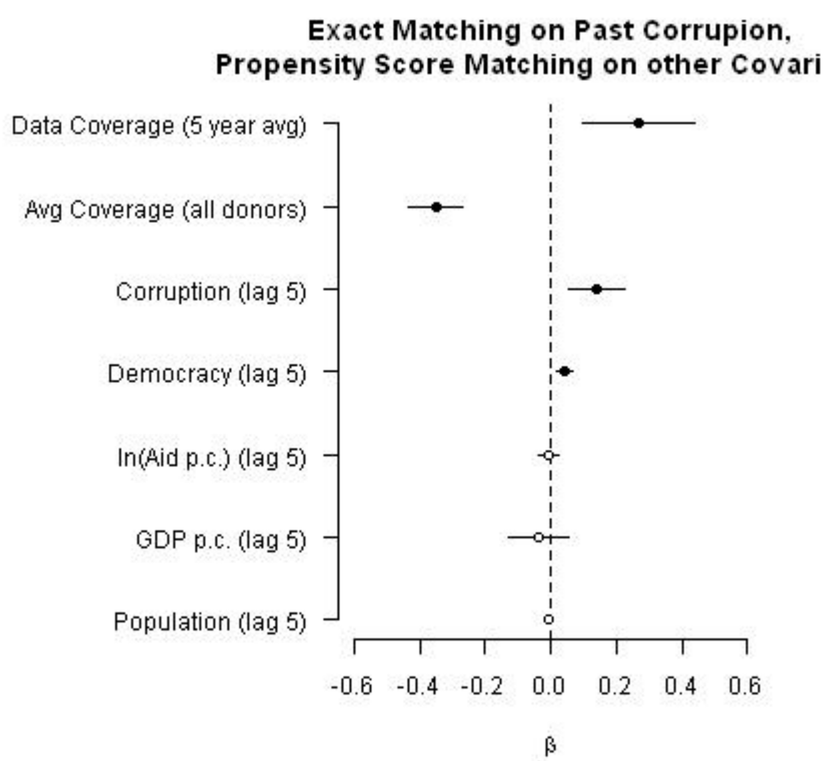

Figure 3: A model of the effect of the past five years of aid transparency on corruption levels in the subsequent three years using a matched dataset that was formed by exact matching on past corruption followed by one-to-one propensity score matching. Black points are statistically significant coefficients (and their 95 percent confidence intervals) and white points are statistically insignificant. Coefficients for some controls such as regional dummies are not reported. [Fix Typo in Title]

Conclusion

We are encouraged by the results, but by no means do we claim the findings are definitive. More work remains to develop different proxies for aid transparency, including the weighting of different AidData fields according to their perceived importance. We will also multiply impute the missing data to help guard against sample bias. It may also make sense to assess the effects of aid transparency for multilateral and bilateral donors separately. Additionally, we plan to select individual cases systematically and qualitatively trace through the historical processes to ascertain if the causal mechanisms we posit actually occurred in the cases. So, as is often the case in such research, there is much more to do.

However, if the findings hold through these future iterations of the project, readers ought to see the results as significant not just statistically but also in substance. While most in the aid community – despite their fierce debates over other questions – have called for greater transparency in foreign aid, much about aid still hides in the shadows of international relations. Even seemingly simple information such as the financial terms of the aid, the implementing agent, and the amount of money disbursed are often left blank in official reports. It is one thing to justifiably decry this lack of information for the non-disclosure it represents. And from a normative perspective, we join our voices with the chorus.

But it is a step beyond these prescriptive calls for action to present robust statistical evidence that failure to transparently report aid leads to increased corruption among recipients. With this paper we have made a strong first move in this direction. And the results, while supporting our initial hypothesis, even surprised us in the potential magnitude of the causal effects. Our overarching point, supported by these preliminary findings, can also be put more hopefully: if donors really want to reduce corruption among their aid recipients, and we have strong reasons to believe this is the case for many donors, then donors themselves may significantly dampen corruption by making their aid more transparent in the first place.

References

Adsera, A., C. Boix, and M. Payne. 2003. Are you being served? Political accountability and quality of government. *Journal of Law Economics & Organization* 19, no. 2: 445-490.

Alchian, Armen A. and Harold Demsetz. 1972. Production, Information Costs, and Economic Organization. *The American Economic Review* 62: 775-95.

Bailard, C. S. 2009. Mobile phone diffusion and corruption in africa. *Political Communication* 26, no. 3: 333-353.

Barder, Owen. 2009. *Beyond Planning: Markets and Networks for Better Aid*. Center for Global Development Working Paper 185.

Barder, Owen. 2006. *Are the Planned Increases in Aid Too Much of a Good Thing?* Center for Global Development Working Paper 90.

Barnett, Michael and Martha Finnemore. 2004. *Rules For The World: International Organizations In Global Politics.* Ithaca, NY: Cornell University Press.

Billger, S. M. and R. K. Goel. 2009. Do existing corruption levels matter in controlling corruption? Cross-country quantile regression estimates. *Journal of Development Economics* 90, no. 2: 299-305.

Bjorkman, M. and J. Svensson. 2009. Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics* 124, no. 2: 735-769.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite. 2003. Is more information better? The effects of "report cards" on health care providers. *Journal of Political Economy* 111, no. 3: 555-588.

Easterly, William. 2006. *The White Man's Burden Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York: Penguin Press.

Easterly, William. 2009. "What Would the Poor Say," Public Talk accessible at *www.nyu.edu/fas/institute/dri/William**Easterly**.ppt*

Eisensee, T. and D. Stromberg. 2007. News droughts, news floods, and us disaster relief. *Quarterly Journal of Economics* 122, no. 2: 693-728.

Ferraz, C. and F. Finan. 2008. Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics* 123, no. 2: 703-745.

Francken, N., B. Minten, and J. F. M. Swinnen. 2009. Media, monitoring, and capture of public funds: Evidence from madagascar. *World Development* 37, no. 1: 242-255.

Gavazza, A. and A. Lizzeri. 2007. The perils of transparency in bureaucracies. *American Economic Review* 97, no. 2: 300-305.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabth A. Stuart. 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*. 15, 3: 199-236.

Kiewiet, D. Roderick, and Mathew D. Mccubbins. 1991. *The Logic of Delegation* Chicago: University Of Chicago Press.

Lyne, Mona M.. *The Voter's Dilemma and Democratic Accountability: Explaining the Democracy-Development Paradox*. University Park: Pennsylvania State University Press, 2008.

Martens, Bertin, Uwe Mummert, Peter Murrell, and Paul Seabright. 2008. *The Institutional Economics of Foreign Aid*. 1 ed. New York: Cambridge University Press.

Mathew D. McCubbins and Thomas Schwartz. 1984. Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28, 1 (Feb.): 165-179.

Moe, Terry. 1990. Political Institutions: The Neglected Side of the Story. *Journal of Law, Economics, and Organization* 6:213-53.

Moyo, Dambisa. 2010. *Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa*. New York: Farrar, Straus and Giroux.

Nielson and Tierney. 2003. Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization* 57 (2): 241-277.

Olken, B. A. 2007. Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy* 115, no. 2: 200-249.

Prat, A. 2005. The wrong kind of transparency. *American Economic Review* 95, no. 3: 862-877.

Ronald Coase. 1937. The Nature of the Firm. *Economica* 4, 16 (Nov.): 386-405.

Sachs, Jeffrey. *The End of Poverty: Economic Possibilities for Our Time*. 2006. Reprint, Boston: Penguin (Non-Classics), 2006.

Swamy, A., S. Knack, Y. Lee, and O. Azfar. 2001. Gender and corruption. *Journal of Development Economics* 64, no. 1: 25-55.

Treisman, D. 2007. What have we learned about the causes of corruption from ten years of cross-national empirical research? *Annual Review of Political Science* 10: 211-244.

Appendix

PLAID Fields Used to calculate project transparency

donor_project_id	additional_info_original_language
crsid financing_agency	Plaid_flow_type
contact	Flow_name
role_of_contact	Flow_code
recipient	number_repayments_per_year
implementing_agency	repay_type
channelcode	Loan_term
channelname	grace_period
borrower	interest_rate
beneficiary	grant_element
commitment_amount	Cancelled
commitment_amount_currency	
commitment_amount_usd_constant	repay_date_first
commitment_amount_usd_nominal	repay_date_last
total_project_cost	untied_amount_usd_nominal
total_project_cost_currency	partial_tied_amount_usd_nominal
total_project_cost_usd_constant	Tied_amount_usd_nominal
total_project_cost_usd_nominal	received_amount_usd_nominal
disbursement_amount	irtc_amount_usd_nominal
disbursement_amount_currency	expert_commitment_amount_usd_nominal
disbursement_amount_usd_constant	export_extended_amount_usd_nominal
disbursement_amount_usd_nominal	export_credit_amount_usd_nominal
title	outstanding_amount_usd_nominal
title_original_language	arrears_principal_amount_usd_nominal
short_description	arrears_interest_amount_usd_nominal
long_description	future_ds_principal_amount_usd_nominal
long_description_original_language	future_ds_interest_amount_usd_nominal
additional_info	interest_amount_usd_nominal

Table A1: A Propensity Model for *Median Coverage (5 year)*

	Estimate	Standard Error
Median Coverage (Year Avg.)	1.55**	0.058
Australian aid (5 year avg)	1.59	2.22
Austrian aid	-3.87	3.59
Belgian aid	-3.74	3.81
Canadian aid	3.05	3.46
Danish aid	-0.510	4.43
Finnish aid	-0.612	9.29
French aid	0.931	1.18
German aid	0.302	0.183
Italian aid	-1.51	1.14
Japanese aid	-0.149	0.171
Dutch aid	-2.98	2.77
New Zealand aid	-671.3**	210.3
Norwegian aid	2.36	5.69
Swedish aid	3.58*	1.73
Swiss aid	-6.33	8.46
US aid	-0.0278	0.220
UK aid	0.595	1.028
Corruption (lag 5)	0.142**	0.036
polity2 (lag 5)	-0.022**	0.0077
Aid p.c. all sources (lag 5)	-0.040**	0.015
CIRI political rights (lag 5)	0.032	0.037
CIRI political rights^2 (lag 5)	-0.0029	0.0035
GDP p.c. (lag 5)	1.57*	0.746
GDP p.c.^2 (lag 5)	-0.092	0.047
ln(Population) (lag 5)	-0.370	0.383
Ln(Population)^2 (lag 5)	0.021	0.020
ln(Aggregate donor trade) (lag 5)	0.109	0.087
ln(Aggregate donor trade)^2 (lag 5)	0.0040	0.0029
Donor Alliance (lag 5)	-0.048*	0.019
UK colony	0.196	0.286
French colony	0.192	0.298
Portuguese colony	0.024	0.332
Spanish colony	0.271	0.325
Dutch colony	0.706	0.655
USSR satellite	0.605	0.581
Belgian colony	-0.457	0.572

Continued on the following page

Continued from previous page

	Estimate	Standard Error
Colony of other	0.162	0.340
Socialist	0.465**	0.166
Cold War (lag 5)	-0.039	0.094
Cold War × Socialist (lag 5)	-0.534**	0.119
War (lag 5)	0.0089	0.078
Sub-saharan Africa	0.296	0.321
Latin America	-0.129	0.327
Middle East/North Africa	0.034	0.352
East Asia/Pacific	-0.393	0.329
Trade with Ireland (lag 5)	-0.026**	0.0096
Trade with Italy (lag 5)	-0.038	0.037
Trade with Luxembourg (lag 5)	-0.041**	0.0083
Trade with Netherlands (lag 5)	-0.177**	0.049
Trade with Norway (lag 5)	0.013	0.0098
Trade with Portugal (lag 5)	0.0080	0.0090
Trade with Spain (lag 5)	-0.0024	0.014
Trade with Sweden (lag 5)	0.0087	0.013
Trade with Switzerland (lag 5)	-0.0043	0.017
Trade with UK (lag 5)	-0.026	0.031
Trade with Japan (lag 5)	0.019	0.022
Trade with New Zealand (lag 5)	-0.016*	0.0066
Constant	-11.9**	3.3
N	1229	
Number of countries	88	
R squared	0.72	

* $p < 0.05$, ** $p < 0.01$. Model includes country-level random intercepts.