Evaluating the impact of the Extractive Industries Transparency Initiative (EITI) on corruption in Zambia

Abstract

The Extractive Industries Transparency Initiative (EITI) is internationally recognised as a leading anti-corruption scheme, which promotes transparency, accountability and good governance of public oil, gas, and mining revenues. The paper provides the first rigorous quantitative investigation of the impact of EITI on corruption in Zambia. Using a case-comparison approach, called the Synthetic Control Method (SCM), we find that the implementation of EITI provoked a significant decrease in corruption in Zambia (with the corruption-reducing effect of EITI being, though, much stronger at the earlier stages of implementation).

Keywords: EITI; Corruption; Transparency; Extractive Industries; Zambia

1. Introduction

The Extractive Industries Transparency Initiative (EITI) has emerged in the international policy arena as a key measure to curb corruption in the extractive sector. Sovacool et al. (2016) explains that the initiative aims to ensure free, full, and independent assessments of how extractive companies interact with government. For states to acquire and maintain EITI membership, they must enforce contract disclosure, adhere to financial transparency standards and audits (also known as *reconciliation*), and ensure the running of a multi-stakeholder group, which provides civil society a platform to keep public officials and companies accountable for their activities in this sector.

In 2016, a surge of interest arose from academic scholars publishing quantitative evaluations of the effect of EITI on corruption; for example see Kasekende et al. (2016), Öge (2016), Papyrakis et al. (2016) and Sovacool et al. (2016). Whilst these evaluations largely focused on providing estimates of the *average* effect of EITI across member countries, a notable point which has emerged from such work is the need to further conduct quantitative research on individual member states (as to develop a better understanding of EITI and how to improve it). Building on these arguments, we contribute to the existing literature on the effect of EITI on corruption by performing the first quantitative case-comparison country analysis on the EITI member state of Zambia.

Zambia is a highly mineral-endowed state, which sits on top of the largest known reserves of copper in Africa. It makes a particularly interesting case-study in this research domain given that the country has been prominently highlighted within the corruption literature for its continued internal battle with political and public corruption. For example, the work of Taylor (2006) has highlighted the impeachment of former Presidents regarding such issues. More broadly, Zambia makes an interesting case-study, given the endemic corruption and overall poor track record in resource management in most parts of Sub-Saharan Africa. Notably, a study released by the African Union estimated that corruption cost African countries $150 billion a year. This is more than 6 times the amount the continent receives in aid from developed economies (Council on Foreign Relations, 2009). Meanwhile, EITI (2010) identifies Africa as a region of specific interest to the initiative given that many countries across the continent face considerable governance issues.

To put things into context, in April 2007, the Zambian government appealed to the World Bank to assist them in administering a scoping study (regarding the relevant costs and benefits

of EITI participation). The corresponding review, in September 2007, suggested that the proposed implementation of EITI would be both feasible and beneficial to the country (World Bank, 2007). Following these recommendations, in 2008, the Zambian government announced its commitment to implementing EITI and during 2009 became an official EITI candidate. Finally, by 2012 (after a successful validation) Zambia was designated a fully compliant member of EITI.[1] To date, the Zambian EITI (ZETI) national body has published seven EITI reports which include reconciliation data. These published reports have gradually expanded coverage on financial transactions in the mining, gas and oil sector, with the number of reporting companies increasing from 16 to 40 (EITI, 2016).

In this study, we aim to investigate the hypothesis that the adoption of EITI has reduced the prevalence of corruption in Zambia, as measured by two popular measures of corruption – i.e., the Transparency International's Corruption Perceptions Index (CPI) and the World Governance Indicators' Control of Corruption Index (CCI). To do so, we implement a case-comparison estimation method, derived from Abadie and Gardeazabal (2003) and Abadie et al. (2010), called the Synthetic Control Method (SCM). Adoption of this comparative methodology, first, offers to mitigate common limitations associated with conventional time-series case study analysis (such as the need to extrapolate over time) and, second, strengthens the selection procedure by choosing a comparison unit to act as a counterfactual to Zambia.

Furthermore, we also investigate whether the beneficial effects of EITI are concentrated in particular stages of its implementation. This builds on the work of Corrigan (2014) and Papyrakis et al. (2016), who suggest that the effects of the intervention may occur prior to a country obtaining compliant status. For example, Papyrakis et al. (2016) claim that the effects of EITI are strongest whilst a country is in its candidate stage (i.e. the initial stage of implementation in preparation for compliance validation). We also build upon the only other known study, by Etter (2014), to have used SCM to measure the effects of EITI. Our analysis extends this earlier work with the inclusion of inference tests to evaluate the significance of reported findings and by concentrating attention to a fully compliant member of the EITI standard.

In the next section of the paper we briefly review existing empirical evidence on the EITI-corruption nexus. Section 3 describes our estimation strategy and data. Section 4 presents the

---

[1] Also see the official EITI standard by the EITI International Secretariat (2016) for further details on the requirements and process of joining the EITI.

results of our empirical investigation. We find that corruption decreased substantially during the earlier stages of EITI implementation (in contrast to the insignificant changes observed in the later candidacy and compliance stages). Finally, Section 5 concludes by summarising our main findings and offering a critical reflection on the Zambian EITI process.

2. The effect of the Extractive Industries Transparency Initiative on corruption

Most recently, a series of cross-country studies have empirically examined the effect of EITI on corruption. Conclusions from such works are largely mixed as evident from the varied spectrum of results. For example, whilst Sovacool et al. (2016) and Öge (2016) report, on average, that EITI has not had a significant effect on corruption, Corrigan (2014) and Papyrakis et al. (2016) find a significant corruption-reducing effect (and Kasekende et al., 2016, a significant corruption-enhancing effect instead). Conflicting results generated a methodological debate, with Kasekende et al. (2016), for example, arguing that other EITI studies do not sufficiently control for selection bias. Two points of particular interest have emerged from this methodological debate.

First, several studies claim that the effect of the EITI on corruption may occur prior to a country obtaining compliance membership status and that the effect may differ across the different stages of implementation. For example, Corrigan (2014) describes that government actions at the onset of EITI implementation (when public authorities express their public commitment to the scheme) can lead to a significant reduction in corruption. Meanwhile, Papyrakis et al. (2016) claim that EITI is more effective in combating corruption at the stage of implementation during which countries receive a candidate status – this is because during this stage candidate countries must implement a series of important changes, including the timely publication of EITI reports and public disclosure of all related financial flows in the extractive sector, in order to be designated as fully compliant. Further empirical investigation would help clarify how the EITI-corruption nexus might differ across different stages of implementation for member states.

Second, Kasekende et al. (2016) and Papyrakis et al. (2016) have highlighted the need to conduct further context-specific investigations (as a means to improve our understanding of where and why EITI may or may not work). McCartney (2006) emphasises that by solely focusing on average effects, important and insightful effect heterogeneity may remain undetected. Isaksson and Ng (2006) also suggest that, from a policymaker's point of view, a

cross-country study does not answer whether the intervention has actually served its purpose in a specific context of interest. However, known examples of studies which seek to identify the direction or magnitude of the effect of EITI on corruption in national contexts are largely limited to a published study by Sovacool and Andrews (2015) and a conference paper by Etter (2014).

Sovacool and Andrews (2015) evaluate the effect of Azerbaijan's and Liberia's compliance with EITI in 2009, by simply depicting the time-trend of the CCI between the years 2006-2012. The authors concluded that whilst trends in corruption improved after compliance in Azerbaijan, this was not the case for Liberia. Further contextual analysis by the authors suggests that a reduction in political support of the initiative may have caused implementation to stall in Liberia (in addition to disputes within the country's EITI multi-stakeholder group).

Meanwhile, Etter (2014) applied two different counterfactual approaches, SCM and Entropy Balancing, to measure the effect of EITI on corruption in Mali and Peru. Both of these methods utilised pre-processing techniques that assign a set of optimised weights to a control group of non-EITI countries in order to formulate a counterfactual. The notable difference between these methods, and arguable advantage of SCM, is that SCM tests the similarity of the constructed weighted control unit across a pre-EITI time period. This helps to distinguish whether unobserved differences may be confounding the estimated weightings applied to the control units. Etter (2014) concludes that, whilst corruption fell in Peru following the introduction of EITI, in Mali corruption increased. His study, though, suffers from methodological limitations by only analysing the period prior to the countries becoming compliant, as well as failing to measure the statistical significance of the SCM results.

3. Evaluating the effect of EITI on perceptions of corruption in Zambia

Building on these insights, we employ the SCM approach in order to examine the effect of EITI on corruption in Zambia. We pursue this case-comparison methodological approach for two reasons. First, SCM mitigates common weaknesses associated with conventional time-series case study analysis, since it limits the need to extrapolate over time and weakens the susceptibility of the analysis to time-varying confounding (Jandoc et al., 2015; Linden and Arbor, 2015). Second, SCM prevents the erroneous task of descriptively choosing a comparison unit, which, if inappropriately chosen, may confound conclusions. Instead of relying on subjective decision-making, SCM employs a data-driven approach to select a

comparison with a statistical affinity with Zambia (Abadie et al. 2010). As an additional methodological innovation, rather than simply assuming that a single comparison unit exists that is sufficiently similar to Zambia, SCM explores whether a weighted combination of potential comparison units provides a better fit. Using notation and description adapted largely from Abadie et al. (2010), Cavallo et al. (2013), and Galiani and Quistorff (2016), this section continues by formally defining the SCM strategy and inference methodology adopted, and further describes the data employed in the analysis.

Synthetic Control Method Estimation Strategy

Let's start by defining an index in the range $\{1, ..., J+1\}$, where the first unit corresponds to Zambia and consecutive ones to the remaining J are non-EITI countries, which may serve as a control unit to Zambia. These control units are also known as 'donors' or the 'donor pool'. Furthermore, T represents the total amount of periods (years) observed, so that $\mathbf{Y}_j$ is defined as the $(T \times 1)$ vector of corruption outcomes for unit J. This vector of corruption outcomes may also be divided into pre and post EITI vectors, so that $\mathbf{Y}_j = \overleftarrow{\mathbf{Y}_J} \backslash \overrightarrow{\mathbf{Y}_j}$. Also, to further disaggregate $\mathbf{Y}_j$, let $\mathbf{Y}_0$ represent the $(T \times J)$ matrix of corruption outcomes for all donors and $\mathbf{Y}_1$ the $(T \times 1)$ matrix of corruption outcomes for Zambia. Therefore, the pre-EITI corruption outcomes can be retrospectively represented by $\overleftarrow{\mathbf{Y}_1}$ and $\overleftarrow{\mathbf{Y}_0}$.

Next, we define $\mathbf{X}$ as a set of k pre-EITI predictors of corruption ($\mathbf{Y}$). Specifically, whilst $\mathbf{X}_0$ represents the $(k \times J)$ matrix of donor predictors, $\mathbf{X}_1$ represents the $(k \times 1)$ matrix of predictors for Zambia. $\mathbf{V}$ is then taken to be a $(k \times k)$ variable weighted matrix, indicating the relative significance of the predictor variables ($\mathbf{X}$) in determining corruption ($\mathbf{Y}$). The final matrix to be defined in this model is defined as $\mathbf{W}$, which is a $(J \times 1)$ weighted matrix $(w_2, w_3, ..., w_{j+1})'$ where $\sum_{j=2}^{J+1} w_j = 1$ and $w_j \geq 0 \ \forall j \in \{2, ..., J+1\}$. Note that a weighted average of the donors' corruption outcomes, or more simply put, the 'synthetic unit' or 'synthetic Zambia', can now simply be given by $\mathbf{Y}_0\mathbf{W}$.

The purpose of the SCM approach is then to identify the optimal combination of weightings in $\mathbf{W}$ that minimises the distance $\sqrt{(\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})} = \ || \ \mathbf{X}_1\text{-}\mathbf{X}_0\mathbf{W} \ ||_\mathbf{V}$. For the purposes of this analysis, the distance is defined as the root mean squared prediction error (RMSPE). To further explain the reasoning behind this method, assume that corruption ($\mathbf{Y}$) is affected by both observed ($\mathbf{X}$) and unobserved ($\mathbf{U}$) factors ($\mathbf{Y} = \beta\mathbf{X} + \mathbf{U}$). The SCM process

searches for weights **W** that result in similar **X** values – when this consecutively leads to a small absolute value of $\overleftarrow{\overline{Y_1}}\text{-}\overleftarrow{\overline{Y_0}}\mathbf{W}$, Abadie et al. (2010) argue that this will also imply similar values for **U**.

In the case that pre-treatment outcome trajectories are well matched, the impact estimate for each post-EITI year can simply be taken as the difference between Zambia's corruption score and the synthetic unit's one. Or more formally:

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j \geq 2} w_j Y_{jt}$$

However, one limitation (with regards to the conventional method of simply taking the differences between the treated unit and the synthetic unit) is that this does not allow for a straightforward analysis of varying treatment effects across time. For example, it would be of interest to examine if the treatment effect is stronger during different stages of EITI implementation. In order to achieve this, the analysis would require to compute the changes in differences for each period. For this reason, to standardise the methodology used, our study calculates the treatment effects as the average annual change in differences:

$$\hat{\alpha}_{1t} = \frac{Y_{1t-t_0}}{t - t_0} - \frac{\sum_{j \geq 2} w_j Y_{jt-t_0}}{t - t_0}$$

Method of Inference

It is important to evaluate whether the differences between the synthetic Zambia and real Zambia are statistically significant. However, as explained by Abadie et al. (2015), the use of conventional inference techniques applied to regression analysis is inappropriate due to the small sample nature of the data. Alternative methods which have been developed to conduct inference tests in this domain relate to permutation tests (Cavallo et al., 2013). Here, the same SCM approach is simply re-applied to the donor pool to create a distribution of placebo study effects (also known as a distribution of 'in-place' placebo effects, see Galiani and Quistorff, 2016). Conceptually, the estimated impact of EITI in Zambia would be undermined if we estimate effects of similar magnitudes in countries that are not members. Ando (2015) explains that this concept is akin to classical randomisation inference, except that the treatment is not randomised. Therefore, the interpretation of this method is informal but nevertheless very informative.

In practice, a p-value is derived from the proportion of non-EITI countries which have an estimated effect at least as large as that of the treated units. Using notation from Galiani and

Quistorff (2016), when the estimated effect for Zambia is represented by $\hat{\alpha}_{1t}$ and the distribution of the corresponding placebos derived from the non-EITI countries is $\hat{\alpha}_{1t}^{P}=\{\hat{\alpha}_{jt}: J \neq 1\}$, then the two sided p-value is given by:

$$\text{p-value} = \Pr(|\hat{\alpha}_{1t}^{P}| \geq |\hat{\alpha}_{1t}|) = \frac{\sum_{j \neq 1} 1(|\hat{\alpha}_{jt}| \geq |\hat{\alpha}_{1t}|)}{J}$$

In addition to the above, two further methods of inference have been proposed by Abadie et al. (2010). Specifically, these methods are more considerate of the quality of the approximations of the synthetic units in the donor pool. Poorly estimated synthetic units can result in overly conservative p-values, because the larger deviation in the estimation error underestimates the relative rarity of the treated unit's effect size (see Galiani and Quistorff, 2016). For this reason, the second method proposed simply conducts the procedures discussed above on a restricted donor pool. This comprises only of donors that can offer a similar pre-EITI RMSPE as Zambia. Abadie et al. (2010) further proposes a number of parameters by which to restrict the donor sample. This includes restricting donors to just those that have a RMSPE that is at most twice as large (RS(2)), 5 times as large (RS(5)) or 20 times as large (RS(20)) as the treatment unit (although the need to endorse a subjective judgement on an appropriate cut-off point poses a clear weakness).

A further weakness of this inference method is that, for it to be informative, it requires that the size of the restricted sample size is sufficient. For example, should one want to obtain a confidence level of 5%, the sample must host at least 20 constituents. Mathematically, the minimum rank percentage value here is computed as $\frac{1}{20} = 5\%$. Consequently, when restricting the sample size of the donor pool, a distinct trade-off is found between improving the fit of the donor pool and reducing the level of confidence which can be inferred. Therefore, we only report p-values for the restrictions discussed above which still enable the analysis to infer confidence at the 5% level or lower.

The third method suggested by Abadie et al. (2010) mitigates the need to neither restrict the donor pool nor choose a subjective cut-off point. To do so, it specifically assumes that the size of the deviation between the real and synthetic units during the pre-intervention period is informative for assessing the deviation in the treatment period. In order to prevent observing overly conservative p-values, this test simply adjusts the estimated effects measured in the treatment period against the pre-EITI deviation. Practically, this involves dividing the post-EITI effect size estimate by the pre-EITI RMSPE. Similar to the first method, p-values are

derived from the proportion of the non-EITI countries which have an estimated effect at least as large as that of the treated unit (Zambia) (Galiani and Quistorff, 2016).

To distinguish between these three methods of inference, we refer to the first method discussed as the Non-Restricted Donor Sample method (NRDS), the second one as the Restricted Sample method (RS(n), where n refers to the cut off parameter), and the third one as the Adjusted Non-Restricted Donor Sample method (ANRDS).

Data

The study period used to investigate the effects of EITI on corruption in Zambia is constrained to the years 2002-2014 (max: T=12). The period is selected to maximise the availability of control countries and predictor variables for the estimation of synthetic Zambia. The pre-EITI period in this analysis is defined as the period 2002-2006, whilst the years 2007-2014 define the treatment period of interest.

Specifically, we define 2007, the year of the World Bank Review, as the first intervention year by extending arguments derived from Corrigan (2014). Here, it is suggested that the government's initial actions, such as a public expression of commitment to EITI, may have signalled a change in the government's tolerance of corrupt activities. By initiating the scoping review through the World Bank, the Zambian government might have wished to provide a credible signal about its intentions to reform and implement anti-corruption measures. As a result, this may have reduced the prevalence of corrupt activity to the extent that corrupt agents sought to protect themselves from future risk of exposure.

With regards to outcome variables, the study adopts two different annual perception measures of corruption, i.e. the CPI and the CCI indices. Transparency International's (2011) CPI index ranks countries based on how corrupt their public sector is perceived to be. The index is scaled from 0 to 10 with higher scores associated with lower levels of corruption. The CCI indicator measures the extent to which public power is exercised for private gain. The index is scaled from -2.5 to 2.5, also with higher scores corresponding to lower levels of corruption (Kaufmann et al., 2009).

Both indicators are aggregate corruption indices, i.e, they standardise and average the scores from many individual sub-indices in order to obtain a single weighted value for each country (for a detailed review of the technicalities relating to the weighting procedures for the CPI and

CCI, see Rohwer, 2009). We focus on these two indicators, given their extensive used as outcome variables in the existing empirical EITI literature to date (which allows for greater comparability to previous findings in this domain).

Figure 1. Perceptions of corruption in Zambia (2002-2014)
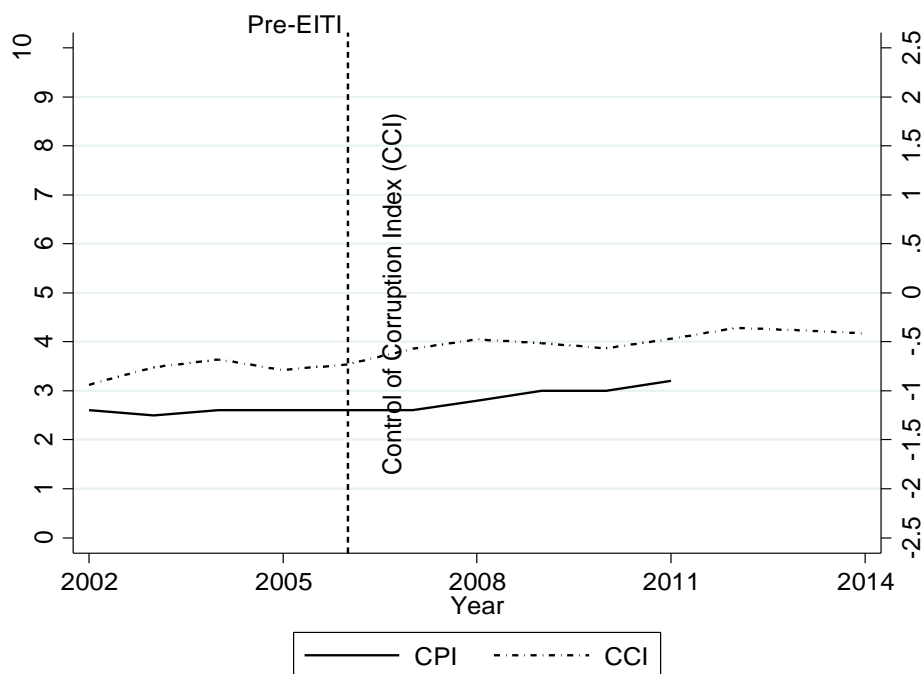
A practical issue to bear in mind is that, since 2012, Transparency International changed the way of measuring its CPI index (and its scale). For this reason, our analysis is only conducted until 2011 on this indicator, since the new CPI index is not comparable to its pre-2012 equivalent. Figure 1 presents the trends of the CCI and CPI outcome indicators before and after 2006 (the baseline year). This shows that the index values for both indicators increased since 2006 (in other words, perceived corruption has decreased since Zambia expressed interest in the EITI scheme).

Naturally, little can be said with regards to how corruption would have evolved in the absence of the EITI, without adopting strong assumptions to extrapolate over time. Therefore, in order to construct a counterfactual (synthetic) Zambia using the SCM, we compile a list of predictor variables of corruption (**X**). The complete list of predictors is presented in Appendix 1. This list comprises of variables that have been identified as significant determinants of corruption (in the existing EITI cross-country empirical literature, as well as the broader literature on corruption and its causes) – for some key studies, see Treisman (2000), Jain (2001), Ali and Isse (2003), Serra (2006), Ata and Arvas (2011) and Elbahnasawy and Revier (2012). Further to this, and in line with other SCM studies, such as Abadie et al. (2015), Sills

et al. (2015) and Ando (2015), we also include the pre-EITI corruption indices as predictor variables. It may well be that unobserved characteristics prove highly influential in determining the corruption outcome variables (which could, then, drive differences between the synthetic and real pre-EITI trajectories in corruption). Therefore, including the pre-intervention outcome variable as a predictor helps to limit the sensitivity of the weighting procedure to these unobservable factors and ensures a better fitting pre-intervention model.

A list of the countries constituting each donor pool, according to each outcome indicator, is available in Appendix 2. Donors included in the restricted sample inference tests, mentioned above, are also indicated in this table. Appendix 3 provides descriptive statistics for the donor pools predictor variables, based on which the Synthetic Zambia is constructed (note, that the predictor variables are taken as the average value of each variable during the pre-EITI period, 2002-2006, in line with common practice discussed in Abadie et al (2010; 2015).

4. Results

After deriving a set of weights for the donor pool based on an optimization algorithm, the synthetic Zambia is constructed as a weighted average of all donor countries. The individual weights derived for each donor country using this optimization tool are reported in Appendix 2 by outcome indicator.[2]

It is customary to assess whether the estimated synthetic unit is an appropriate counterfactual to the real Zambia. In order to do so, SCM analyses, such as the ones by Abadie et al. (2010) and Ando (2015), visually compare the difference between the pre-treatment outcomes of the real unit and the synthetic units. This assesses whether the differences in unit predictors (see Appendix 4) and unobserved variables significantly affect the synthetic units ability to predict the treated unit over time. Following this method, Figure 2 presents the corruption outcome trends over the study period for both Zambia and the Synthetic Zambia. When observing the trends in corruption across the outcome indicators until 2006 (CPI – top panel; CCI – bottom panel), it is evident that the synthetic units are very similar to Zambia. The levels and changes in corruption over this period are almost identical in both outcome graphs.

---

[2] The optimisation algorithm used in this analysis is available through the Stata SCM command –synth–.See: https://web.stanford.edu/~jhain/synthpage.html for further information.

Figure 2. Comparison of Zambia and Synthetic Zambia's trends in corruption
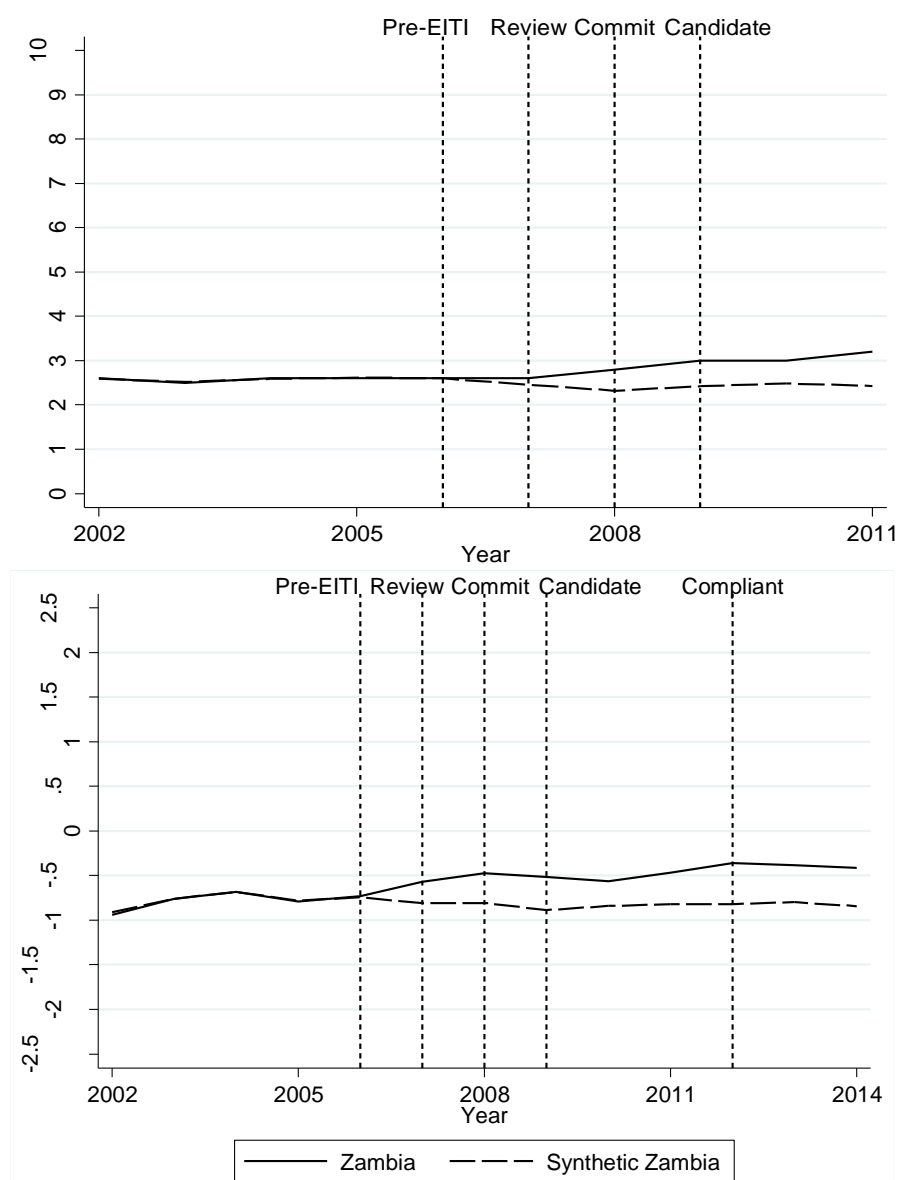
Having concluded that the pre-EITI trends are similar, we now review the estimated effects of EITI from 2007 onwards. Beginning by simply observing the changes over time in Figure 2, it is evident from both the CPI and CCI graphs that Zambia's outcome positively diverges from Synthetic Zambia's values, starting from the World Bank 'Review stage' in 2007. This positive divergence (signalling a drop in corruption) persists throughout all subsequent stages of implementation (i.e. also in 2008 when Zambia officially expressed commitment to the EITI scheme – 'Commit stage' – in 2009, when it received candidate status – 'Candidate stage' – as well as since 2011, when it was declared fully compliant – 'Compliant stage'). For a clearer image of the size of the divergence over time (and across different stages of implementation) also see the long dashed line in Figure 3. This line depicts the difference, or the *gap* as Abadie et al. (2010) describe it, between Zambia and Synthetic Zambia's corruption outcomes in each

time period. Overall the results suggest that EITI has had a beneficial effect in reducing corruption in Zambia.

Figure 3. 'In Place' placebo gaps
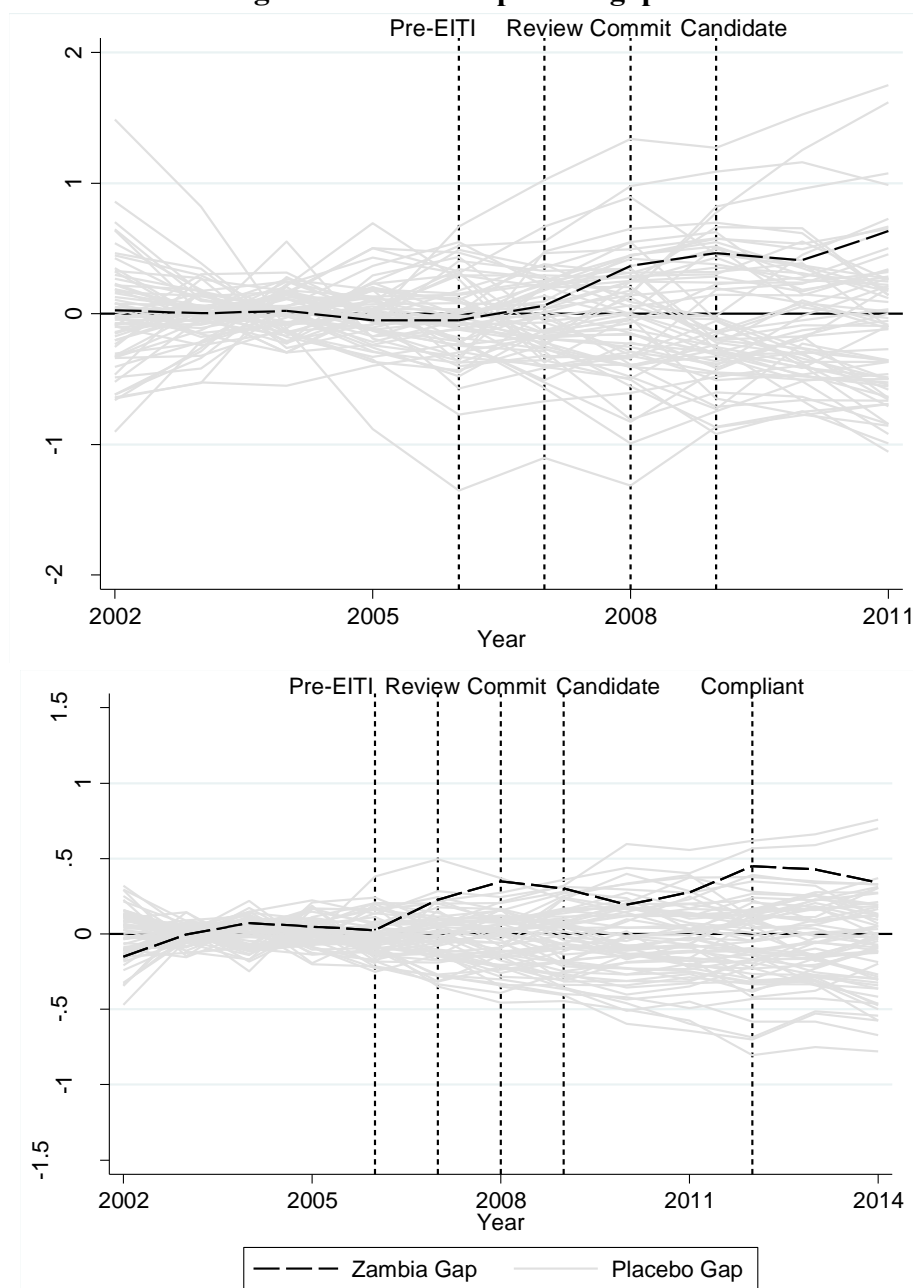
Table 2 provides the overall (average annual) change in gap size for each outcome indicator during the defined treatment (post-EITI) period, as well as the change observed during each individual EITI implementation stage. The latter aim to highlight the stages of the EITI process that resulted in the largest statistically-significant changes in corruption. In order to infer whether these results could have been achieved by chance, we develop a distribution of 'in place' placebo gaps. A graphical representation of the estimated placebo gaps compared to Zambia's can be seen in Figure 3. Table 2 provides the p-values (statistical significance) for

each gap estimate based on the three inference methods that are derived from the placebo gap distributions.

Table 2. Zambia's average annual change in gap size

		Total 2007-14†	Review 2007	Commit 2008	Candidate 2009-11	Compliant 2012-14
CPI	Δ Gap Size (annual)	0.13	0.11	0.31	0.09	*n/a*
	NRDS P-value	0.26	0.39	0.83	0.55	*n/a*
	RS(20) P-value	0.16	0.60	0.12	0.52	*n/a*
	ANRDS P-value	0.02**	0.05**	0.02**	0.15	*n/a*
CCI	Δ Gap Size (annual)	0.04	0.20	0.12	-0.02	0.02
	NRDS P-value	0.32	0.10*	0.26	0.71	0.78
	RS(2) P-value	0.50	0.05**	0.10*	0.95	0.95
	ANRDS P-value	0.23	0.04**	0.06*	0.92	0.97

Note: † CPI results are average annual changes until 2011; *, **correspond to a 10 and 5% level of significance.

The CPI results are the first set of findings presented in Table 2. Here it is shown that the 0.13 average annual increase in Zambia's CPI (i.e. fall in corruption) relative to Synthetic Zambia for the period 2007-11 is statistically significant at the 5% confidence level according to the ANRDS method of inference, but insignificant according to the NRDS and RS(n) criteria. In order to further explore possible causes for this initial disagreement between inference methods, the CPI graph in Figure 3 is re-examined.

Figure 3 (CPI graph) shows that some of the synthetic units developed for the donor countries do not reproduce the average donor (gap) outcomes well (i.e. the distance between the average donor (gap) dotted line and the individual synthetic unit grey lines in the pre-EITI period, is large). For example, one synthetic unit starts almost 1.5 index points higher than the average donor gap outcome in 2002. Given the poor fit of some of the donor pool synthetic

units, it can be argued that perhaps the first method of inference (NRDS) is not particularly suitable and that the donor sample may need to be restricted or adjusted. However, the CPI model did not host many placebo effects where individual synthetic units provided a good fit for Zambia. Therefore, the strictest parameter of those previously discussed, that could be applied, is the loosest restriction of RS(20). The concern now lies with the issue that the RS(n) method may be generating an overly conservative estimate of the size of the Zambian gap given the wide breadth of estimation error this restriction parameter allows. This, therefore, offers some explanation as to why the CPI's RS(20) p-values may disagree with the p-value obtained from the ANRDS method.

Given these limitations regarding the use of the RS(n) and NRDS methods, the ANRDS inference results are deemed the most reliable for inferring the significance of the CPI-model effects (with the 0.13% average annual increase in Zambia's CPI index, during 2007-14, being statistically significant at the 5% level). Further analysis of the estimated effects observed for individual stages of implementation shows that the relative increases during the review (+0.31) and commitment (+0.11) years are also significant at the 5% level (however, the average annual change during the later candidate stage (+0.09) is insignificant according to all three inference tests).

The second set of results presented in Table 2 refer to the CCI model. For the CCI index, the inference tests unanimously agree that the total average annual change in differences (+0.04 index points) is statistically insignificant. However, further investigation of the estimated changes during the individual stages of implementation reveals that there are statistically significant reductions in corruption during the report and commitment years (in line with the earlier findings of the CPI models). Here, a relative increase of 0.20 index points is observed during the earlier review stage (and is statistically significant according to all inference methods; significance at the 10% level according to NRDS and at the 5% level when using our preferred RS(n) and ANRDS p-values). At the same time, the RS(n) and ANRDS tests indicate that the relative increase of 0.12 points during the commitment period is significant at the 10% level. The average annual changes in differences during the later candidacy and compliance phases are almost zero (-0.02 and 0.02) and insignificant.

Overall, it is interesting to observe that the most significant and largest relative changes in both sets of results are witnessed prior to either the candidacy or compliance stage. This is in line with theory presented in Corrigan (2014) which suggests that initial actions, that credibly

signal future implementation of EITI, are likely to dissuade corrupt activities (with corrupt actors reacting proactively to the incoming changes enforced by EITI). In this scenario, it is possible that EITI's largest effects are, in effect, observed much before the full compliance stage. These results also contradict previous findings from Papyrakis et al. (2016), who find that the candidacy stage has the most pronounced corruption-reducing effect.

It may also be the case that practical challenges stifled progress at later stages of implementation. For example, Zambia's EITI reconciliation reports express concerns about independent organisations (acting on behalf of ZEITI) having no legal authority to enforce compliance at the company level (BDO, 2015). Whilst the international EITI secretariat enables member countries to join without enacting its standards into law, the lack of legal authority can hamper the complete auditing of the sector and create systemic corruption loopholes. In addition, there is very limited information (see Zambia EITI, 2015) highlighting how discrepancies identified in reconciliation reports are resolved – a more transparent follow-up procedure with publically available documentation could increase scrutiny of such cases and further deter corruption. Last, the role of multi-stakeholder group meetings can be strengthened (in the anti-corruption arena as well as in other areas) by standardising the reporting of issues raised at these events and setting a formal follow-up process that addresses them.

5. Conclusion

Since its inception, the EITI has been widely recognised by the international community as a major anti-corruption scheme that promotes transparency, accountability and good governance in the extractive sector. Yet, whilst several recent studies have offered quantitative empirical evidence on its effect on corruption, their investigation has been largely confined to cross-country analysis with little attention given to country-specific effects and particularities. Our study contributes to this empirical literature by concentrating attention to the specific corruption-reducing effect of EITI for Zambia. In addition, we also examine whether the effect of EITI is concentrated in particular stages of engagement with the EITI process. Our empirical analysis suggests that EITI reduced corruption in Zambia, especially during the earlier stages of implementation. In later phases of implementation (candidacy and compliance stages), the EITI effect, though, is statistically insignificant and/or of a small magnitude.

This may be associated with practical challenges that possibly hamper EITI's corruption-reducing effect at later stages of implementation. For example, there is currently no independent legal authority in Zambia that can enforce compliance at the company level and, hence, ensure a complete auditing of the extractive sector. In addition, there is also no formal follow-up procedure that tackles discrepancies identified in reconciliation reports and takes forward recommendations from multi-stakeholder group meetings. In that respect, the EITI secretariat could incentivise compliant countries to take further steps in increasing transparency in the sector by introducing a tiered membership system (e.g. one that rewards more proactive states with more advanced membership status).

Our study is a first step in exploring the EITI-corruption nexus in the Zambian context. Zambia was only designated a fully compliant EITI member in 2012 and some of the scheme's corruption-reducing effect may only appear with some time lag. Future research could, hence, revisit our analysis and examine whether the effect of the intervention may change after years of longer experience with the initiative. Another direction of future research would involve replicating our SCM analysis in other contexts, as a means to shed additional light on the specific experience of individual EITI members with the scheme.

References

Abadie, A., Diamond, A. and Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490), pp.493–505.

Abadie, A., Diamond, A. and Hainmueller, J., 2015. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), pp.495–510.

Abadie, A. and Gardeazabal, J., 2003. The economic costs of conflict: A case study of the Basque Country. *The American Economic Review*, 43(1), pp.113–132.

Ali, A.M. and Isse, H.S., 2003. Determinants of Economic Corruption: A Cross-Country Comparison. *Cato Journal*, 22(3), pp.449–466.

Ando, M., 2015. Dreams of urbanization: Quantitative case studies on the local impacts of nuclear power facilities using the synthetic control method. *Journal of Urban Economics*, 85, pp.68–85.

Ata, A. and Arvas, M., 2011. Determinants of economic corruption: a cross-country data analysis. *International Journal of Business and Social Science*, 2(13), pp.161–169.

BDO, 2015. Seventh report of the Zambia Extractive Industries Initative (ZEITI). Available at: http://www.zambiaeiti.org/index.php/ct-menu-item-13 (accessed 08.08.2016).

Cavallo, E., Galiani, S., Noy, I. and Pantano, J., 2013. Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, 95(5), pp.1549–1561.

Corrigan, C., 2014. Breaking the resource curse: Transparency in the natural resource sector and the extractive industries transparency initiative. *Resources Policy*, 41(1), pp.17–30.

Council on Foreign Relations, 2009. Corruption in Sub-Saharan Africa. Available at: http://www.cfr.org/africa-sub-saharan/corruption-sub-saharan-africa/p19984 (accessed 08.08.2016).

EITI, 2016. Homepage Extractive Industries Transparency Initiative. Available at: https://eiti.org/ (accessed 08.08.2016).

EITI, 2010. Impact of EITI in Africa Stories from the ground. *The EITI International Secretariat*. [Online]. Available at: https://eiti.org/files/EITI Impact in Africa.pdf (accessed 08.08.2016).

EITI International Secretariat, 2016. The EITI Standard 2016. Available at: https://eiti.org/node/4922 (accessed 08.08.2016).

Elbahnasawy, N. and Revier, C., 2012. The determinants of corruption: Cross-country-panel-data analysis. *The Developing Economies*, 50(4), pp.311–333.

Etter, L., 2012. Can Transparency Reduce Corruption? Evidence from Firms in Peru and Mali on the Impact of the Extractive Industries Transparency Initiative (EITI) on Corruption. Presented at: The Doing Business Conference at Georgetown University. Washington DC, February 20-21. Available at: *https://repository.library.georgetown.edu/handle/10822/557839* (accessed 08.08.2016).

Freedom House, 2016. *Freedom of the Press Scores and Status Data 1980-2016*. Available at: https://freedomhouse.org/report-types/freedom-press (accessed 08.08.2016).

Galiani, S. and Quistorff, B., 2016. The synth_runner Package: Utilities to Automate Synthetic Control Estimation Using. *University of Maryland*. Available at: econweb.umd.edu/~galiani/files/synth_runner.pdf (accessed 08.08.2016).

Isaksson, A. and Ng, T.H., 2006. Determinants of productivity: cross-country analysis and country case studies. *United Nations Industrial Development Organization Research and Statistics Branch Staff Working Paper*, 1. Available at: www.unido.org/fileadmin/import/63219_determinants_of_productivity_country.pdf (accessed 08.08.2016).

Jain, A.K., 2001. Corruption: A Review. *Journal of Economic Surveys*, 15(1), pp.71–121.

Jandoc, R., Burden, A. and Mamdani, M., 2015. Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations. *Journal of clinical, 68, pp.950*–956.

Kasekende, E., Abuka, C. and Sarr, M., 2016. Extractive industries and corruption: Investigating the effectiveness of EITI as a scrutiny mechanism. *Resources Policy*, 48, pp.117–128.

Kaufmann, D., Kraay, A. and Mastruzzi, M., 2009. Governance matters VIII: aggregate and individual governance indicators, 1996-2008. *World Bank Policy Research Working Paper 4978*. Available at: http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1424591. (accessed 08.08.2016).

Kaufmann, D., Kraay, A. and Mastruzzi, M., 2006. Measuring corruption: myths and realities. *Development Outreach, The World Bank*. Available at: http://www1.worldbank.org/publicsector/anticorrupt/corecourse2007/Myths.pdf (accessed 08.08.2016).

Kolstad, I. and Wiig, A., 2009. Is Transparency the Key to Reducing Corruption in Resource-Rich Countries? *World Development*, 37(3), pp.521–532.

Linden, A. and Arbor, A., 2015. Conducting interrupted time-series analysis for single-and multiple-group comparisons. *Stata Journal*, 15(2), pp.480–500.

McCartney, M., 2006. Can a heterodox economist use cross-country growth regressions?

Post-autistic Economics Review, 37(4), pp.45–54.

Öge, K., 2016. Which transparency matters? Compliance with anti-corruption efforts in extractive industries. *Resources Policy*, 49, pp.41–50.

Papyrakis, E., Rieger, M. and Gilberthorpe, E., 2016. Corruption and the Extractive Industries Transparency Initiative. *The Journal of Development Studies*, In Press..

PRS Group, 2016. *International Country Risk Guide Researchers Dataset*. Available at: http://epub.prsgroup.com/index.php/products/icrg/researcher-dataset-icrg-t3b-political-risk (accessed 08.08.2016).

Rohwer, A., 2009. Measuring Corruption: A comparison between the transparency international's corruption perceptions index and the world bank's worldwide governance. *CESifo DICE Report*. Available at: http://econpapers.repec.org/article/cesifodic/v_3a7_3ay_3a2009_3ai_3a3_3ap_3a42-52.htm (accessed 08.08.2016).

Rydland, L., Arnesen, S. and Østensen, Å., 2007. Contextual data for the european social survey: An overview and assessment of extant resources. *Norwegian Social Science Data Services*. Available at: http://www.nsd.uib.no/om/rapport/nsd_rapport124.pdf. (accessed 08.08.2016).

Serra, D., 2006. Empirical determinants of corruption: A sensitivity analysis. *Public Choice*, 126, pp.225–256.

Sills, E., Herrera, D., Kirkpatrick, A.J., Brandão, A. Jr., Dickson, R., Hall, S., Pattanayak, S., Shoch, D., Vedoveto, M., Young, L. and Pfaff, A., 2015. Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. *PloS one*, 10(7). Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132590 (accessed 08.08.2016)

Solt, F., 2016. The Standardized World Income Inequality Database. SWIID Version 5.1. *Social Science Quarterly*, In Press.

Sovacool, B. and Andrews, N., 2015. Does transparency matter? Evaluating the governance

impacts of the Extractive Industries Transparency Initiative (EITI) in Azerbaijan and Liberia. *Resources Policy*, 45, pp.183–192.

Sovacool, B., Walter, G., Van De Graaf, T. and Andrews, N., 2016. Energy Governance, Transnational Rules, and the Resource Curse: Exploring the Effectiveness of the Extractive Industries Transparency Initiative (EITI). *World Development*, 83, pp.179–192.

Taylor, S.D., 2006. Divergent politico-legal responses to past presidential corruption in Zambia and Kenya: catching the "big fish", or letting him off the hook? *Third World Quarterly*, 27(2), pp.281–301.

Transparency International, 2011. Corruption Perceptions Index In detail. Available at: http://www.transparency.org/cpi2011/in_detail (accessed 08.08.2016).

Transparency International, 2016. *Corruptions Perceptions Index Dataset*. Available at: http://www.transparency.org/research/cpi/overview (accessed 08.08.2016).

Treisman, D., 2000. The causes of corruption: a cross-national study. *Journal of Public Economics*, 76(3), pp.399–457.

World Bank, 2007. Extractive Industries Transparency Initiative (EITI) scoping study for the Republic of Zambia. *World Bank Group Report No. 82644*. Available at: http://documents.worldbank.org/curated/en/134691468166493476/Extractive-Industries-Transparency-Initiative-EITI-scoping-study-for-the-Republic-of-Zambia (accessed 08.08.2016).

World Bank, 2016. World Development Indicators (WDI) Online. Available at: http://data.worldbank.org/data-catalog/world-development-indicators (accessed 08.08.2016).

Zambia EITI, 2015. 2015 Zambia EITI Annual Activity Progress Report. Available at: http://www.zambiaeiti.org/index.php/ct-menu-item-13 (accessed 08.08.2016).

Appendices

Appendix 1. Table of predictor variables used to estimate Synthetic Zambia

Variable Name	Variable Description	Data Source
Freedom of the Press	Degree of print, broadcast, and digital media freedom. Index Range: 0-100 (Lower scores given to nations with a freer press).	Freedom House (2016)
Log of GDP per capita	Log of Gross Domestic Product (GDP) Per Capita, PPP (constant 2011 international US$).	World Bank (2016)
GDP Growth	Annual GDP growth %.	World Bank (2016)
Urbanization	Urban population (% of total).	World Bank (2016)
Labour Participation	Labour force participation rate, total (% of total population ages 15+).	World Bank (2016)
Openness to foreign trade	Trade (% of GDP).	World Bank (2016)
Log of population	Log of population size, total.	World Bank (2016)
Natural Resource Rents	Total Natural Resources Rents (% of GDP).	World Bank (2016)
Inequality	Gini coefficient of income inequality. Index Range: between 0 and 100 (larger values correspond to more unequal income distributions).	Solt (2016)
Bureaucratic quality	Institutional strength and quality of government services. Index Range: 0-4 (higher scores given to nations with higher institutional strength and quality).	PRS Group (2016)
Democratic Accountability	Index of how responsive a government is to its people. Index Range: 0-6 (higher scores are given to nations considered more democratically accountable).	PRS Group (2016)
Ethnic Tensions	Degree of tension within a country attributable to racial, nationality, or language divisions. Index Range: 0-6 (higher scores given to nations where tensions are considered low).	PRS Group (2016)
Religious Tensions	Measure of religious tensions arising from the domination of governance by a single religious group Index Range: 0-6 (higher scores given to nations where tensions are considered low).	PRS Group (2016)
Government Stability	Ability of a government to stay in office. Index Range: 0-12 (higher scores given to nations with governments deemed more stable).	PRS Group (2016)
Internal Conflict	Level of political violence in a country. Index Range: 0-12 (higher scores given to nations with less conflict).	PRS Group (2016)
External Conflict	Risk of foreign action, such as cross-border conflict and war. Index Range: 0-12 (higher scores given to nations with less conflict).	PRS Group (2016)
Law and Order	Strength and impartiality of the legal system and the degree of popular observance of the law. Index Range: 0-6 (higher scores given to nations with more impartial legal systems with lower crime rates).	PRS Group (2016)
Military in Politics	Degree of military participation in politics. Index Range: 0-6 (lower scores given to nations with less military participate in politics).	PRS Group (2016)
Pre-EITI Outcome Variables	CPI or CCI index, dependent on the outcome indicator being modelled.	Transparency International (2016) / World Bank (2016)

Note: Predictors variables are averages (2002-2006).

Appendix 2. List of donor pool countries and weights by corruption outcome indicator

CPI				CCI			
Country	**Weight**	**Country**	**Weight**	**Country**	**Weight**	**Country**	**Weight**
Australia	0*	Malaysia	0*	Armenia	0.004*	Lithuania	0
Austria	0	Mexico	0*	Australia	0	Luxembourg	0
Bangladesh	0	Moldova	0	Austria	0*	Malaysia	0*
Belarus	0	Morocco	0	Bangladesh	0	Malta	0
Belgium	0	Netherlands	0*	Belarus	0*	Mexico	0*
Bolivia	0	New Zealand	0*	Belgium	0	Moldova	0
Brazil	0	Nicaragua	0.355*	Bolivia	0	Morocco	0
Bulgaria	0	Pakistan	0	Brazil	0	Netherlands	0*
Canada	0	Panama	0	Bulgaria	0	New Zealand	0*
Chile	0*	Paraguay	0	Canada	0	Nicaragua	0
China	0.032*	Poland	0	Chile	0	Pakistan	0
Costa Rica	0	Portugal	0	China	0	Panama	0
Croatia	0	Romania	0	Costa Rica	0	Paraguay	0
Czech Republic	0	Russia	0.006*	Croatia	0	Poland	0*
Denmark	0	Singapore	0*	Czech Republic	0*	Portugal	0
Ecuador	0*	Slovak Republic	0	Denmark	0*	Romania	0*
Egypt	0*	Slovenia	0*	Ecuador	0.421	Russia	0
El Salvador	0	South Africa	0	Egypt	0	Singapore	0
Estonia	0	Spain	0	El Salvador	0	Slovak Republic	0
Finland	0*	Sri Lanka	0	Estonia	0	Slovenia	0
France	0	Sweden	0*	Finland	0*	South Africa	0*
Greece	0*	Switzerland	0	France	0	Spain	0
Hungary	0*	Thailand	0*	Greece	0	Sri Lanka	0*
Iceland	0*	Tunisia	0	Hungary	0*	Sweden	0*
India	0	Turkey	0.054	Iceland	0	Switzerland	0*
Ireland	0	Uganda	0*	India	0.141	Thailand	0
Israel	0	Uruguay	0.161	Iran	0	Tunisia	0
Italy	0	Venezuela	0*	Ireland	0	Turkey	0.139
Japan	0	Vietnam	0*	Israel	0	Uganda	0*
Jordan	0	Zimbabwe	0.239	Italy	0*	Uruguay	0
Kenya	0*			Japan	0	Venezuela	0*
Korea Rep.	0.154*			Jordan	0	Vietnam	0
Latvia	0			Kenya	0	Zimbabwe	0*
Lithuania	0			Korea Rep.	0.296		
Luxembourg	0			Latvia	0		

Note: * indicates that the country was included in the restricted donor pool sample during inference tests referred to as RS(n).

Appendix 3. Descriptive statistics (Zambia and donor pools)

Variable	Zambia	CPI Donor Pool			CCI Donor Pool		
		Mean	Min	Max	Mean	Min	Max
Log of population	16.27	16.53	12.56	20.99	16.47	12.56	20.99
Log of GDP per capita	7.78	9.56	7.03	11.41	9.55	7.03	11.41
Freedom of the press	63.8	37.45	8	90	38.17	8	90
Inequality	50.43	36.22	22.07	22.07	36.18	22.07	57.52
Labour Participation	79.92	60.59	40.2	86.7	60.23	40.2	86.7
Urbanization	36.19	64.98	12.25	100	65.42	12.25	100
Openness to foreign trade	63.61	89.34	21.16	430.35	89.64	21.16	430.35
GDP Growth	6.72	4.47	-16.99	18.28	4.58	-16.99	18.28
Natural Resource Rents	11.39	4.68	0	43.54	5.03	0	46.06
Religious tensions	4.6	4.83	1	6	4.79	1	6
Military in politics	5	4.45	0	6	4.47	0	6
Law and order	4	4.20	0.5	6	4.19	0.5	6
Internal conflict	9.44	9.83	2.96	12	9.81	2.96	12
Government stability	6.28	8.67	5.04	11.5	8.66	5.04	11.5
External Conflict	10.05	10.39	6.13	12	10.34	6.13	12
Ethnic tensions	5	4.21	1	6	4.25	1	6
Democratic Acc.	4.25	4.68	1	6	4.65	1	6
Bureaucratic quality	1	2.66	1	4	2.63	1	4
CPI	*2.58*	5.05	1.2	9.7	-	-	-
CCI	*-0.77*	-	-	-	0.40	-1.48	2.55

Appendix 4. Comparison of predictor variable balance between

Zambia and Synthetic Zambia

Variable	Zambia	CPI Synthetic Zambia	CCI Synthetic Zambia
Log of population	16.27	16.38	17.42
Log of GDP per capita	7.78	8.17	8.22
Freedom of the press	63.8	60.36	47.88
Inequality	50.43	46.21	46.24
Labor	79.92	69.52	68.88
Urbanization	36.19	49.32	38.31
Openness to foreign trade	63.61	64.13	49.98
GDP Growth	6.72	1.18	5.53
Natural Resource Rents	11.39	10.70	11.23
Religious tensions	4.6	4.20	3.68
Military in politics	5	2.43	2.63
Law and order	4	2.88	2.93
Internal conflict	9.44	8.25	8.02
Government stability	6.28	7.45	7.29
External Conflict	10.05	9.37	10.08
Ethnic tensions	5	4.15	3.00
Democratic Acc.	4.25	4.06	4.19
Bureaucratic quality	1	1.46	2.13
Average CPI	*2.58*	2.58	-
Average CCI	*-0.77*	-	-0.77