Electoral Systems and Corruption

Julie K. Faller,† Adam N. Glynn,‡ and Nahomi Ichino§

Abstract

What is the effect of electoral systems on corruption? Persson, Tabellini and Trebbi (2003) proposed that plurality electoral systems should lead to lower corruption compared to proportional representation (PR) systems because the former creates a direct link between voters and politicians whom voters can hold accountable for corruption. The empirical question remains unresolved, however, in part due to the endogeneity of the electoral institutions and difficulties in measuring corruption. Using nonparametric methods and new data to reduce sensitivity to these problems, we find no evidence for this hypothesis. Instead, we find some evidence in the opposite direction, that PR leads to less corruption.
1 Introduction

What is the effect of electoral systems on corruption? A Petition to Restore Power to the People recently circulated in the Republic of South Africa called for electoral system reform to create a closer connection between voters and their elected representatives in order to help prevent “public servants [... from] using their position to enrich themselves and their families” (Agang South Africa 2013). This followed a 2009 report from the Independent Panel of Assessment of Parliament, which argued that South Africa’s party list system makes Members of Parliament more accountable to party leaders than to the electorate (Independent Panel Assessment of Parliament 2009). Adding a “constituency-based element into the electoral system” to connect representatives more closely to voters is now part of the platform of the opposition party Democratic Alliance (Democratic Alliance 2013).

Corruption – defined as the misuse of public office for private gain (Treisman 2000) – is an urgent problem in South Africa and around the world (Transparency International 2012). It is linked to greater income inequality (Gupta et al. 2002), lower levels of socially-productive innovation (Murphy et al. 1993), inefficient investment by firms (Choi and Thum 2004, Svensson 2003), and lower human capital (Reinikka and Svensson 2005). Corruption is also associated with lower trust in and satisfaction with the political system (Anderson and Tverdova 2003, Chang and Chu 2006, Seligson 2002) and lower voter turnout (Davis et al. 2004, McCann and Domínguez 1998). Although reformers cannot easily change the cultural and historical-institutional legacies associated with corruption (Serra 2006, Treisman 2000, Treisman 2007), changes to the electoral system might dislodge this political-economic equilibrium of low public engagement and high corruption.

The logic behind this call for electoral reform in South Africa echoes the dominant scholarly view that citizens want to vote out corrupt politicians and that political institutions that empower
citizens to do so will lead to lower corruption. Persson et al. (2003, 961-2) contends that plurality systems should lead to lower corruption than proportional representation (PR) systems, because the former creates a direct link between voters and politicians whom voters can hold accountable for corruption. In the same framework, Kunicová and Rose-Ackerman (2005) argues that while plurality systems offer more opportunities for corruption to rank-and-file politicians, PR systems offer those opportunities to party leaders. Because voters are less able to monitor and sanction party leaders in a PR system, corruption will be greater in PR systems. Treisman (2007) challenges the robustness of both results, and there is currently little consensus on how electoral systems affect the extent of politicians’ misuse of public office for private gain.

Several difficulties leave this question unresolved. First, corruption is difficult to measure (Apaza 2009, Arndt and Oman 2006) and the available indices of perceptions of corruption measure slightly different concepts. Scholars have been rightly cautious about their use. Second, the endogeneity of a country’s electoral system to its politics is generally not sufficiently addressed. Although cross-sectional analyses of corruption often include controls for factors that likely affect the choice of electoral systems (Treisman 2007), the calculations of strategic actors highlighted in the theoretical and case-specific literature on electoral system choice are varied and difficult to capture. Consequently, the estimated causal effects may be biased, and the direction of this bias is difficult to characterize. Third, cross-sectional studies generally overlook when an electoral system was adopted. Some control variables are therefore measured pre-treatment for some units and post-treatment for other units, and the resulting estimated average effect is a weighted average of different types of treatment effects. Persson et al. (2003) uses panel data and fixed effects to address some of these problems, but Treisman (2007) questions the reliability of the corruption measure was available to Persson et al. (2003).

We address these difficulties in several ways. First, as in a fixed effects analysis, we focus on within-country changes from one electoral system to another. By noting the timing of the treatment, we know when the confounding variables should be measured. Requiring information
on the pre-transition electoral system also effectively limits the analysis to countries that did not simultaneously transition to democracy and change various other political institutions. This leaves us with ten countries that adopted closed list PR systems between 1998 and 2009, and we adopt a Fisherian randomization inference framework for the statistical analysis to accommodate this small number of observations. Second, to ameliorate confounding issues, we use pre-treatment values of corruption to construct a synthetic control unit for each of our treated units (Abadie et al. 2010). A country that changes its electoral system is compared to a weighted average of countries that have the same pre-transition electoral system and are in the same region of the world. Third, to address some of the concerns with measurement of the outcome variable, we use the nonparametric signed rank statistic to test for statistical significance as discussed in Glynn and Ichino (2014). These data issues and methods are described in greater detail in the following sections.

We find no support for the hypothesis that PR leads to more corruption. In fact, we find some evidence suggesting that PR might lead to less corruption, although this evidence is not statistically significant at traditional levels. Two robustness checks further strengthen our results. First, we improve the matches with a differences-in-differences approach and obtain similar results. Second, we introduce a dose-weighted version of the signed rank statistic (Rosenbaum 2002, 2009) in order to account for the heterogeneity in the pre-transition electoral systems. Despite using several different approaches, including the fixed effects regressions mirroring Persson et al. (2003)’s specifications (online appendix), we were unable to find support for the findings of Persson et al. (2003).

There are a number of important caveats to our results. First, they are limited in scope to the countries for which we observed the adoption of closed list PR. Persson et al. (2003)’s hypothesis may still hold for countries out of this sample. Second, not enough time may have passed for effects to accrue in these ten countries. Third, our analyses use measures of corruption based upon expert surveys. We have used a nonparametric approach so as to be less sensitive to measurement problems, and the results are relatively stable across different indices. But a different measure of
corruption that relies less on perceptions may still reverse these findings. Finally, the synthetic control and difference-in-differences approaches may not have eliminated all bias due to unmeasured confounding, which is likely if the measures of corruption we use are poor. A Rosenbaum (2002) style sensitivity analysis on these results is relatively straightforward. However, this would only weaken our confidence that PR might be reducing corruption for these countries, not strengthen the evidence for the Persson et al. (2003) hypothesis.

Despite these caveats, our results suggest that further work is necessary before the Persson et al. (2003) hypothesis should be accepted or used to guide institutional design. Either new measures of corruption need to be developed or extant theories should be re-examined for what purposes and under what conditions voters hold politicians accountable for corruption. In particular, theory more appropriate for conditions in new democracies can contribute to the scholarship on corruption.

In the following section, we discuss the existing theoretical perspectives on how electoral systems affect corruption. We discuss three broad difficulties with current empirical analyses of this question and describe a methodological approach to address these issues, before presenting our results.

2 Electoral Rules, Accountability, and Corruption

Electoral rules for legislative bodies define how votes are converted into a set of legislators following an election. These rules may be very complex, but the most basic distinction is between PR systems and plurality systems. In PR systems, legislative seats are allocated to political parties on the basis of the total votes won by each party. Voters may express preferences over particular candidates within a party in an open list PR system. But in a closed list PR system, party leaders determine the order in which individual politicians are ranked on the party list. Once the total number of seats awarded to a party is determined, that number of candidates from the top of the list are elected. By contrast, in plurality or majoritarian systems, the candidate or party with the greatest
number of votes wins all the seats in a district. These districts will often only have one seat, and in some cases, the candidate or party must also meet a majoritarian vote threshold such as 50%. As we describe in more detail in Section 5.2, a mixed system uses both of these systems for elections to the same legislative body. While we focus on this broad distinction between plurality and PR systems, electoral systems can also vary in the number of total seats in the legislature, number of districts, and district magnitudes, as well as the exact formula for allocating seats within these broad categories.

Persson et al. (2003, 961–2) proposes that electoral systems affect corruption by changing the incentives for politicians to engage in corruption. Since party leaders determine the placement of politicians on party lists, political parties intervene in the chain of delegation from voters to politicians in PR systems and incentivize politicians to exert effort in and for the party rather than in office for the electorate. Furthermore, because the number of seats won by a party in PR systems depends upon the total number of votes won by the party’s candidates, neither the negative electoral consequences of engaging in corruption nor the electoral rewards from refraining from corruption are internalized by an individual politician as fully as they would be in plurality systems. Consequently, it expects PR to lead to more corruption than do plurality systems, and mixed systems to lead to a middle level of corruption, less than with a PR system but more than with a plurality system.

With a similar logic, Kunicová and Rose-Ackerman (2005) argues that corruption should be worse under closed list than open list PR systems, because political rents are primarily extracted by party leaders rather than by rank-and-file politicians, and voters are less able to hold these party leaders accountable in the former than in the latter. Gingerich (2009) also associates closed list PR systems with more corruption than open list PR systems, proposing that party leaders in closed list PR systems can entice bureaucrats to engage in corruption by promising rewards of positions within the party, a power party leaders lack in open list PR systems.

However, Golden and Chang (2001) and Chang (2005) note that open list PR and plurality
systems could lead to more, not less, corruption than closed list PR, since the former allow voters to favor or disfavor individual politicians. This gives incentives to politicians to “cultivate the personal vote” (Carey and Shugart 1995) and to turn to corruption to finance these activities, with greater pressure where politicians face greater competition such as in high district magnitude systems (Mainwaring 1991, Reed 1994). This contrasts with Myerson (1993) and Persson et al. (2003) which argue that electoral systems with lower barriers to entry and offer more alternative options to voters, like PR systems which usually have higher district magnitude than plurality systems, create more competitive pressures on incumbents to refrain from corruption.

In statistical analyses of cross-national data, Persson et al. (2003), Kunicová and Rose-Ackerman (2005), and Tavits (2007) have found that plurality systems with personal ballots have the lowest levels of corruption, whereas open list PR systems have more corruption than plurality systems, and closed list PR systems have the highest levels of corruption. However, the electoral system is not among the factors that Serra (2006) finds to be consistently associated with corruption. Similarly, Treisman (2007) notes that this effect of electoral system on corruption appears fragile. These results were not robust to the inclusion of control variables or the use of data from different years than the original analyses (232). In the next section, we describe several general difficulties with assessing the effect of electoral systems on corruption in cross-national analyses before discussing our approach to ameliorating these issues.

3 Challenges for Empirical Analysis

There are three major issues for empirical analysis. The first is an imprecise definition of treatment effects due to electoral systems being adopted at different points in time. The second is unmeasured confounding due to non-random treatment assignment. The third is the difficulty in measuring corruption. Each can contribute to the inconsistency and fragility of empirical assessments of the causal effects of electoral systems on corruption.
3.1 Timing of electoral system adoption and post-treatment bias

Most empirical studies of the effects of electoral systems on corruption, including the main analysis in Persson et al. (2003), are cross-sectional analyses with a recent measure of perceived corruption as the outcome ($Y$), an indicator of some aspect of the electoral system as the treatment ($T$), and a set of control variables such as degree of democracy, economic development and freedom of the press ($X$), that previous studies suggest also affect the outcome. One problem with this setup is that the countries in the analysis adopted their electoral systems at different points in time. Some are well-established democracies that have maintained the same type of electoral system for many years (Group 1), like the United States, while others are countries that have recently made a transition to democracy or adopted significant electoral reforms (Group 2), such as the former Soviet republics. By analyzing these two groups together, these studies estimate an average of the long-run effects for Group 1 countries and short-run effects for Group 2 countries.

Moreover, if control variables $X$ are measured at the same point in time for all countries, they are likely measured before the adoption of the electoral system (treatment) for Group 2, but after adoption for Group 1 (Figure 1). Many of these covariates are likely affected by the electoral system $T$, so that the effect of electoral systems on corruption is likely to suffer from post-treatment bias and to be underestimated for Group 1 countries. Under favorable circumstances, the effect for this set of countries may be interpreted as a controlled direct effect, while the effect for Group 2 can be interpreted as a total effect (VanderWeele and Vansteelandt 2009). However, the assumptions needed for the estimation of controlled direct effects are different from the assumptions needed for total effects, and it is unlikely that both sets of assumptions will simultaneously hold.

3.2 Unmeasured confounding from non-randomized electoral systems

The second issue is that electoral systems are not randomly assigned, but originate in a political process. For example, colonial heritage and history of communist rule affect what electoral reforms
Assignment of $T$, Group 1

Assignment of $T$, Group 2

Measurement of $Y$

Measurement of $X$

Short-Term Total Effect

Long-Term Controlled Direct Effect

Figure 1: Coefficient estimates are an average of long-term controlled direct effects and short-term total effects of electoral systems on corruption.

are considered viable options for a country (Blais and Massicotte 1997, Golder and Wantchekon 2004, Luong 2000, Mozaffar 1998). Divided societies might also prefer proportional systems to majoritarian ones (Lijphart 1969), and the number of social cleavages and the size of the country also affect some electoral system choices (Benoit 2007). Analyses of corruption generally include these variables that plausibly affect both the choice of electoral systems and corruption as control variables (Persson et al. 2003, Treisman 2007, Gerring and Thacker 2004).

However, studies of the origins of electoral systems emphasize the politicians’ assessments of their future success under alternative electoral systems. The balance of power among these actors and the uncertainty about future power are crucial to understanding the electoral institutions they choose (Andrews and Jackman 2005, Benoit 2004, Luong 2000, Remington and Smith 1996). Because systematic and reliable information on this uncertainty and politicians’ calculations are difficult to obtain, substantial unmeasured confounding may lead to biased estimates.

Persson et al. (2003) addresses this issue with a fixed effects analysis with data from Political Risk Services’s International Country Risk Guide (ICRG) as the outcome, the only measure with enough coverage at the time of its writing. But there are sudden, unexplained changes in the ICRG scores for particular countries and inconsistencies in the scoring across countries (Treisman 2007, 221), and it appears to be less reliable than alternative measures. Imai and Kim (2013) has also recently shown that country-year fixed-effects estimators rely on questionable implicit
comparisons.

### 3.3 Measurement of corruption

As noted earlier, the enduring challenge to the empirical study of corruption is its definition and measurement. While most scholars agree that corruption is “the misuse of public office for private gain,” they disagree on what constitutes misuse and how to measure it (Della Porta and Vannucci 2012, Philp 1997). Those engaged in illicit activities have incentives to hide corruption, compounding the difficulty of measuring corruption.

Most cross-national empirical studies use two indices of corruption: Transparency International’s Corruption Perception Index (CPI) and the World Bank’s Worldwide Governance Indicator for “control of corruption” (Treisman 2007). Both indices aggregate information from several sources, including expert surveys and business group risk reports. Scholars have registered various concerns with these indicators: they measure perceptions of corruption, not corruption itself; treat expert opinions as independent although experts may rely on the same sources; encourage cross-country comparisons although the relationship between experienced corruption and reported corruption could be heterogenous across countries; and have changed their methodologies thus making inter-temporal comparisons difficult (Apaza 2009, Arndt and Oman 2006). More objective indicators may be available for specific countries. For example, Gagliarducci et al. (2011), Golden and Chang (2001), and Golden and Picci (2005) have used legislator absenteeism, formal complaints against legislators, and an index measuring the gap between funds spent on infrastructure and its quality to measure corruption in Italy. But such measures may indicate different things in different contexts and are difficult to obtain, so they are not well-suited for cross-national comparative studies.

---

2The latter variable is sometimes referred to as “GRAFT.”
4 Definitions, Methods, and Data

We take several steps to address these challenges in our analysis. We use the World Bank’s Control of Corruption (CCE) index like most cross-national empirical studies of corruption, but we use nonparametric methods and rank-based statistics that are less sensitive to measurement errors in the outcome variable, which we describe in greater detail in the next section. We also define treatment to be applied when a country changes its electoral system, similar in spirit to Persson et al. (2003)’s fixed effects analysis. This allows us to use synthetic control methods (Abadie et al. 2011), which are more transparent about what countries are being compared with each other than a fixed effects analysis. Later we take a differences-in-differences approach to ameliorate confounding and estimate short-run total effects for the treated countries.

First, we simplify our analysis by following Persson et al. (2003) and classify electoral systems as (a) proportional representation (PR), which may be open list or closed list; (b) plurality or majoritarian (plurality, henceforth); or (c) mixed. We define treatment as being applied when there is a change from one of the latter systems to closed list PR, and investigate the effect of this new electoral system on corruption for nominally democratic countries that adopted closed list PR. Data availability directed our analysis to changes to, rather than from, closed list PR and to focus on closed list PR rather than open list PR. We date the treatment to the adoption of these rules rather than when a new legislature is elected and seated under the new rules, because we expect forward-looking politicians to decide whether to engage in or refrain from corruption by considering their consequences for their prospects for election under the new rules.

We use three variables from the World Bank’s Database of Political Institutions compiled as part of the Quality of Governance (QoG) time-series dataset (Teorell et al. 2011, 6 April 2011 version). The first is a dummy variable for plurality, which takes the value of 1 if plurality is used to select any member of any chamber of the national legislature or if there is competition for the seats in a one-party state, and 0 otherwise. The second is a dummy variable for PR, which
takes the value of 1 if proportional representation is used to select any member of any chamber of the national legislature, and 0 otherwise. The third is a dummy variable for closed list, which is defined only when the PR variable is 1. We define a country as having a plurality system in a given year when plurality is coded 1 and PR is coded 0 and as having a PR system in a given year when plurality is coded 0 and PR is coded 1. When both variables take the value 1, we consider the country to have a mixed system.

Second, we initially examine the difference between plurality and closed list PR, the comparison for which Persson et al. (2003) expects the largest effect, and then supplement the analysis by looking at the difference in corruption between mixed systems and closed list PR. Because of data limitations on the outcome variable described below, we investigate changes in electoral systems between the years 1998 and 2009. A country must have information on both the pre-transition and new electoral systems to be included in the analysis. This effectively excludes countries that transition to democracy and adopt a variety of new political institutions along with a new electoral system. This is in keeping with the cross-national literature that excludes non-democracies and allows us to distinguish the change in electoral system from other changes to political institutions that might affect corruption. In this time period, Kazakhstan, Kyrgyzstan, Mongolia, and Togo changed from plurality to closed list PR systems. We also find changes in Algeria, El Salvador, Macedonia, Niger, Russia, and Ukraine from mixed systems to closed list PR systems.

Third, having defined treatment and identified the treated countries, we must find appropriate control countries because electoral systems were not randomly assigned. We use the Synth package v. 1.1-3 (Abadie et al. 2011) in R v.2.15.3 to construct a synthetic control unit that is as similar as possible to each of the treated units on pre-treatment values of the outcome variable (corruption) for the three years preceding the electoral system change. Under assumptions that are generally weaker than the assumptions required for linear regression or fixed effects, this approach

Kyrgyzstan is an exception. Because it transitioned to a plurality electoral system two years before transitioning to closed list PR, we include only two years preceding its transition in the construction of its synthetic control.
ameliorates some of the problems due to non-random treatment assignment (Abadie et al. 2010).

The synthetic control is a weighted average of control units drawn from a donor pool of countries in the same region as the treated country that had and maintained the treated country’s pre-reform electoral system. This restriction to the same region helps account for factors such as shared colonial heritage, culture, and diffusion of policies that affect both electoral system choice and corruption. We use the alternative region coding \(ht\_region2\) from Teorell and Hadenius (2005), available from the QoG dataset. This coding groups Mongolia with Eastern Europe due to its post-Communist legacy.

The potential donor countries for Kazakhstan, Kyrgyzstan, Mongolia, and Togo are countries in the same region as the treated country that had plurality systems from 1996 through 2011. The potential donor countries for Algeria, El Salvador, Macedonia, Russia, and Ukraine are countries in the same region as the treated country that maintained mixed closed list systems from 1996 to 2011. Niger used open list in its PR tier before the change in its electoral system, but very few African countries used and maintained this system. We prioritize matching on pre-treatment values of the outcome variable and allow the donor pool for Niger to be composed of countries which maintained any mixed system from 1996 to 2011. Niger is consequently matched to Senegal, which had a mixed system with closed list in its PR tier. Only Madagascar and Tunisia among the potential donor countries were unused (received weight of 0) in the synthetic controls. The countries used as synthetic controls and their relative weights are presented in Tables 1 and 2.

As noted above, our primary measure of corruption is the World Bank Worldwide Governance Indicator for Control of Corruption (CCE) (Kaufman et al. 2013).\(^4\) This indicator measures perceptions of the extent of use of public office for private gain, including petty corruption, grand corruption, and state “capture.” It aggregates 30 data sources, comprising expert assessments from governmental, commercial and non-governmental organizations and surveys of citizens (Kaufman

\(^4\)The World Bank corrected errors in its 2011 estimates for CCE in February 2013 (Kraay 2013). We use the corrected version. QoG provides this World Bank data series through 2009. We obtained the 2010 and 2011 data directly from the World Bank website.
<table>
<thead>
<tr>
<th>Treated Unit</th>
<th>Kazakhstan</th>
<th>Kyrgyzstan</th>
<th>Mongolia</th>
<th>Togo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic Control</td>
<td>1 Tajikistan</td>
<td>0.200 Azerbaijan 0.190 Tajikistan 0.409 Turkmenistan 0.201 Uzbekistan</td>
<td>1 Tajikistan</td>
<td>0.027 Botswana 0.063 Cote d’Ivoire 0.050 Gabon 0.038 Gambia 0.047 Ghana 0.052 Kenya 0.042 Malawi 0.032 Mali 0.047 Mauritius 0.048 Uganda 0.046 Zambia 0.508 Zimbabwe</td>
</tr>
</tbody>
</table>

Table 1: Synthetic Controls using the World Bank’s Control of Corruption Index (CCE) for countries that changed from plurality to closed list PR

<table>
<thead>
<tr>
<th>Treated Unit</th>
<th>Algeria</th>
<th>El Salvador</th>
<th>Macedonia</th>
<th>Niger</th>
<th>Russia</th>
<th>Ukraine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic Control</td>
<td>1 Senegal</td>
<td>0.236 Bolivia 0.166 Guatemala 0.178 Honduras 0.420 Mexico</td>
<td>0.179 Albania 0.167 Armenia 0.406 Croatia 0.128 Georgia 0.121 Lithuania</td>
<td>1 Senegal</td>
<td>1 Albania</td>
<td>0.977 Albania 0.012 Armenia 0.005 Croatia 0.003 Georgia 0.003 Lithuania</td>
</tr>
</tbody>
</table>

Table 2: Synthetic Controls using CCE for countries that changed from mixed systems to closed list PR

et al. 2010). The variable CCE ranges from -2.5 to 2.5 with higher values indicating less perceived corruption. It was measured biennially from 1996 until 2002 and annually thereafter. For years without data, we take an average of the prior and following years’ scores. Because we need the outcome variable measured both before and after treatment, we only examine countries that adopted closed list PR between 1998 and 2009. Tables 3 and 4 present summaries of balance on pre-treatment values of the outcome variable. In all cases, the average pre-treatment values of the synthetic control are closer to those of the treated units than those of the unweighted regional sample.

We prefer this CCE measure to Transparency International’s Corruption Perceptions Index
Treated Unit Average Pre-Treatment CCE for:

<table>
<thead>
<tr>
<th>Country</th>
<th>Treated Unit</th>
<th>Synthetic Control</th>
<th>Regional Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazakhstan</td>
<td>-0.89</td>
<td>-0.97</td>
<td>-1.07</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>-1.10</td>
<td>-1.10</td>
<td>-1.05</td>
</tr>
<tr>
<td>Mongolia</td>
<td>-0.62</td>
<td>-0.97</td>
<td>-1.09</td>
</tr>
<tr>
<td>Togo</td>
<td>-0.98</td>
<td>-0.98</td>
<td>-0.54</td>
</tr>
</tbody>
</table>

Table 3: Average Pre-Treatment CCE for Countries with Plurality to Closed List PR Transitions

<table>
<thead>
<tr>
<th>Country</th>
<th>Treated Unit</th>
<th>Synthetic Control</th>
<th>Regional Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>-0.61</td>
<td>-0.38</td>
<td>-0.13</td>
</tr>
<tr>
<td>El Salvador</td>
<td>-0.62</td>
<td>-0.62</td>
<td>-0.69</td>
</tr>
<tr>
<td>Macedonia</td>
<td>-0.62</td>
<td>-0.62</td>
<td>-0.41</td>
</tr>
<tr>
<td>Niger</td>
<td>-0.88</td>
<td>-0.34</td>
<td>-0.11</td>
</tr>
<tr>
<td>Russia</td>
<td>-0.76</td>
<td>-0.74</td>
<td>-0.29</td>
</tr>
<tr>
<td>Ukraine</td>
<td>-0.68</td>
<td>-0.68</td>
<td>-0.26</td>
</tr>
</tbody>
</table>

Table 4: Average Pre-Treatment CCE for Countries with Mixed Systems to Closed List PR Transitions

(CPI), which similarly takes business group and country expert surveys and creates a 0-10 score with higher scores indicating less corruption (Transparency International 2010). While CPI and CCE use similar data sources, CPI differs from CCE in that scores are standardized based on each country’s percentile rank within the larger sample of countries. CPI is a relative score by construction and was not designed to be compared across years. CCE also covers more countries for more years than CPI\(^5\).

In order to assess the effects of the treatment, we take the difference in CCE scores between the treated and control countries, averaged over time. Because the donor pool is relatively small for each treated country, we cannot use the inferential approach presented in Abadie, Diamond and Hainmueller (2010). Instead we treat each country that adopted closed list PR and its synthetic control as exchangeable – i.e., we treat the synthetic control as if it were a country and assume that treatment was as-if randomly assigned between the treated country and its synthetic control.

\(^5\)Neither measure was available for enough years for Persson et al. (2003) to use in its fixed effects analysis.
This allows us to conduct Fisherian randomization inference (Rosenbaum 2002, Ch. 2) to calculate \( p \)-values without assuming a parametric model or assuming that these countries were sampled from some larger population. If the synthetic controls are sufficiently comparable to the treated units, then the exchangeability assumption holds. This allows us to evaluate the sharp null hypothesis of no effect of electoral system for any country with pairwise randomization inference. In this case, sufficiently comparable means that each synthetic control acts as if it were matched to each treated unit in a pairwise randomized experiment. This assumption is strong, but generally weaker than the assumptions required for causal inference with linear regression or fixed effects that are currently employed in cross-national analyses of corruption.

5 Analysis

With this synthetic control analysis, we find no support for Persson et al. (2003)’s hypothesis that the move to closed list PR from plurality increases corruption. In fact, the analysis suggests that either there is no effect, or even the opposite – that this change leads to lower corruption. Additional difference-in-difference and dose-weighted analyses for robustness only strengthen these results.

5.1 Plurality Systems to Closed List PR

Figure 2 presents our analysis for the four countries that changed from plurality systems to closed list PR systems. The vertical line marks the change in the electoral system for each country. The dark solid line is the country’s score on the World Bank’s control of corruption (CCE) measure, with higher values indicating less corruption. The dotted line is the CCE score of the synthetic control for each country. The dashed line is CCE for the diff-in-diff adjusted synthetic control. Specifically, if there is a difference between the treated country and the synthetic control in the average pre-treatment outcome values, we adjust the synthetic control by the average pre-treatment
difference in order to remove this discrepancy. For Kyrgyzstan and Togo, there is no difference between the dotted line and the dashed line. This indicates good fit for the synthetic control for these countries.

The estimated effect using the Synth package for each treated country is the average vertical distance between the solid and the dotted lines after treatment. These average differences are reported in Table 5 along with the number of years of data that make up these averages.

<table>
<thead>
<tr>
<th></th>
<th>Synth Estimate</th>
<th>Post-Treatment Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazakhstan</td>
<td>0.17</td>
<td>3</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>0.14</td>
<td>3</td>
</tr>
<tr>
<td>Mongolia</td>
<td>0.47</td>
<td>2</td>
</tr>
<tr>
<td>Togo</td>
<td>-0.07</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5: Synth estimates for countries that move from plurality/majoritarian systems to closed list PR

If the move to closed list PR from plurality were to increase corruption as Persson et al. (2003) expects, then we would expect the estimates in Table 5 to be negative, and for the dark solid lines in Figure 2 to be below the dashed line after the point of treatment. Only Togo fits this pattern and has the smallest difference with its synthetic control of these four countries. The other three countries indicate positive effects of plurality on CCE, although these effects are also small. Even when the synthetic controls for Kazakhstan and Mongolia are difference-adjusted so as to better match the pre-treatment values of these countries, the estimated effects are still contrary to the Persson et al. (2003) hypothesis (Table 6). The results in Figure 2 and in Tables 5 and 6 are fairly strong evidence against the Persson et al. (2003) hypothesis for these countries.

6 Mongolia moved to closed list PR from a block vote (multiple non-transferable vote) system, which is an extremely disproportional form of plurality system. Under the previous system, fairly small changes in vote share led to large changes in the composition of the legislature (Schafferer 2005), which Persson et al. (2003) note as a condition that should strongly incentivize politicians to refrain from corruption. It may be that the block vote created a collective action problem for incumbent legislators that weakened this incentive, but we do not find any support for the hypothesis that PR increases corruption.

7 We find similar results using Transparency International’s Corruption Perception Index (CPI) instead of CCE (online appendix).
Figure 2: Control of Corruption (WB CCE) over time, for each country that moved from plurality/majoritarian systems to closed list PR. The solid line is the treated country, the dotted line is the synthetic control, and the dashed line is the synthetic control adjusted for the average pre-treatment difference in CCE scores.
<table>
<thead>
<tr>
<th>Country</th>
<th>Diff-in-Diff Estimate</th>
<th>Post-Treatment Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazakhstan</td>
<td>0.09</td>
<td>3</td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>0.14</td>
<td>3</td>
</tr>
<tr>
<td>Mongolia</td>
<td>0.12</td>
<td>2</td>
</tr>
<tr>
<td>Togo</td>
<td>-0.07</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6: Diff-in-Diff Synth estimates for countries that move from plurality/majoritarian systems to closed list PR

5.2 Mixed Systems to Closed List PR

We conduct the same synthetic control analysis for the six countries that changed from mixed electoral systems to closed list PR systems in this period. Mixed systems (or mixed member systems) are those in which both plurality/majoritarian and PR rules are used for election to the national legislature, or where there is more than one chamber, used for elections to the lower chamber (Massicotte and Blais 1999). While Persson et al. (2003) highlights the proportion of seats that are elected by one formula or the other, scholarship on political outcomes in mixed systems has tended to focus on whether and how the seats in the different tiers are linked. It considers whether the different rules apply in geographically distinct areas so a voter votes either in a PR system or in a plurality system, or whether legislators elected in different systems represent the same geographical area (Moser and Scheiner 2012). Because our primary purpose is to compare PR and plurality systems, we treat mixed systems as falling in between these two “pure” systems.

If we restrict the analysis to this set of countries, there appears to be mild but insignificant support for the Persson et al. (2003) argument (Figure 3 and Table 7). Algeria, Niger, Russia, and Ukraine have lower CCE than their synthetic controls after adopting closed list PR, and Russia has the largest difference in average CCE scores for any pair. The other two countries that moved from

---

8 See Nishikawa and Herron (2004) for a review of other definitions.
9 Unlike the other treated countries which started out using closed list for its PR seats within a mixed electoral system, Niger moved from a mixed system using open list for its PR seats to “pure” PR using closed lists. Open list systems allow voters to express their preferences over individual politicians, not just parties, so they are more like plurality than closed list systems. Therefore Persson et al. (2003)’s argument may be more likely to hold for Niger than the other countries that start out with mixed systems with closed list in its PR tier. However, once we difference-adjust the synthetic control, we find that CCE is greater for Niger than its synthetic control (Table 7).
a mixed system to closed list PR, El Salvador and Macedonia, had average CCE scores greater than their synthetic controls.

<table>
<thead>
<tr>
<th>Country</th>
<th>Synth Estimate</th>
<th>Diff-in-Diff Estimate</th>
<th>Post-Treatment Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Algeria</td>
<td>-0.21</td>
<td>0.02</td>
<td>13</td>
</tr>
<tr>
<td>2 El Salvador</td>
<td>0.26</td>
<td>0.26</td>
<td>13</td>
</tr>
<tr>
<td>3 Macedonia</td>
<td>0.07</td>
<td>0.07</td>
<td>8</td>
</tr>
<tr>
<td>4 Niger</td>
<td>-0.12</td>
<td>0.41</td>
<td>4</td>
</tr>
<tr>
<td>5 Russian</td>
<td>-0.53</td>
<td>-0.52</td>
<td>5</td>
</tr>
<tr>
<td>6 Ukraine</td>
<td>-0.42</td>
<td>-0.42</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 7: Synth and Diff-in-Diff Synth estimates for countries that move from mixed systems to closed list PR

However, there are two complications with the analysis for these countries. First, the synthetic controls for Algeria and Niger do not provide good matches for the pre-treatment outcome values of these countries. Once these synthetic controls are difference-adjusted as in the previous section, the estimated effect reverses sign, as reflected in the solid lines that are mostly above the dashed lines in Figure 3. Second, the move from a mixed system to closed list PR is arguably a smaller “dose” of the treatment than the move from a plurality system to closed list PR. Therefore, these six treated countries should likely be given less weight than the four countries considered in Section 5.1. These issues tend to provide evidence against the Persson et al. (2003) hypothesis. We formally incorporate these points in the combined analysis of the next section.

5.3 Combined Analysis

The analyses above suggest that moving to closed list PR systems may actually reduce corruption. To assess the strength of this evidence, we combine the evidence from the plurality and mixed system countries and present nonparametric tests based on weighted and unweighted signed rank statistics. Although we prefer this approach to standard fixed effects models, we note that all of the substantive points here are robust to the use of standard fixed effects models.

We start with the analysis that is most sympathetic to the Persson et al. (2003) hypothesis. The
Figure 3: Control of Corruption (WB CCE) over time, for each country that moved from mixed systems to closed list PR. The solid line is the treated country, the dotted line is the synthetic control, and the dashed line is the synthetic control adjusted for the average pre-treatment difference in CCE scores.
combined mean of the Synth estimates in Tables 5 and 7 is -0.024 and the median is 0.0, so the overall evidence does not support the Persson et al. (2003) hypothesis. We use the exchangeability assumption discussed earlier to conduct a nonparametric signed rank test, which allows us to assess the strength of this evidence with this small sample size without resorting to parametric assumptions. The Wilcoxon signed rank statistic assigns the ranks 1 (smallest) through 10 (largest) to each of the 10 pairs according to the size of absolute difference in post-treatment CCE scores (averaged over time) between the paired treated country and synthetic control countries. The statistic then sums the ranks only for the pairs where the post-treatment CCE scores are larger for the treated country than for the synthetic control (averaged over time).

This process is presented in Table 8, and the signed rank statistic calculated from our data is 27. If all of the ranks had been positive, it would have been 55. This statistic is difficult to interpret directly, but we can generate a p-value under the sharp null hypothesis of “no effect,” by provisionally assuming that electoral system is irrelevant for corruption and permuting treatment status between the treated countries and their synthetic controls within pairs. Because we have 10 pairs, there are $2^{10} = 1024$ different within-pair permutations of treatment status, of which our observed data represents just one possibility. For each permutation, we re-calculate the signed rank statistic, giving us a permutation distribution of 1024 statistics. Figure 4 presents this null distribution based on the analyses in Tables 5 and 7. The p-value is then determined by calculating the proportion of these 1024 statistics that are more extreme than our observed statistic of 27. This procedure is a nonparametric version of a paired t test and is explained in detail in Glynn and Ichino (2014). As noted above, the evidence is not supportive of the Persson et al. (2003) hypothesis. The two-sided p-value is effectively 1 because as we can see from Figure 4, the observed statistic is in the middle of the permutation distribution.

The mean of the difference-adjusted synth estimates is 0.01 and the median is 0.08 – provisional evidence that closed list PR systems may reduce corruption. A nonparametric signed rank test now produces a two-sided p-value of 0.56. The evidence against the Persson et al. (2003)
Table 8: Signed rank statistic for the combined groups. The absolute value of Macedonia’s estimate is slightly larger than that of Togo, so we include an additional decimal place in the analysis to avoid ties in the rank.

The hypothesis is strengthened further if we consider that theory suggests larger effects for transitions from plurality to closed list PR systems than from mixed to closed list PR systems. The outcomes from transitions from plurality systems might therefore be weighted more heavily than the outcomes from transitions from mixed systems. If we knew the appropriate doses (weights) to assign to these two types of transitions, then we could use the dose-weighted signed rank test described in Rosenbaum (2002, 2009). For example, if the move from plurality to closed list PR represented a dose of 2, and the move from a mixed system to closed list PR represented a dose of 1, then we would alter the signed rank statistic by multiplying the ranks of the plurality to closed list PR estimates by 2 and the ranks of the mixed to closed list PR estimates by one. The procedure for generating the $p$-value is otherwise the same as the permutation procedure described above.

Unfortunately, we do not know the doses to assign to the two types of transitions. However, for any arrangement such that the plurality to closed list PR transitions are assigned a dose at least as large as the dose assigned to the mixed to closed list PR transitions, the $p$-value will be at least as small as those reported above. With doses of 2 and 1 as described above, the $p$-value would be approximately 0.70 for the synth estimates and 0.375 for the difference-adjusted estimates. With
Figure 4: *Permutation distribution for the signed rank statistic based on unweighted Synth estimates*

doses of 3 and 1 as described above, the $p$-value would be approximately 0.625 for the synth estimates and 0.275 for the difference-adjusted estimates. Therefore, depending on how we assess the relative strengths of these two types of transitions, the evidence against the Persson et al. (2003) hypothesis can be quite strong.

### 5.4 Discussion

There are a number of important caveats to these results. First, these results are limited in scope to the countries for which we have changes in electoral systems. Methodologically, it makes sense to limit the analysis to these ten countries where the evidence will be most straightforward to assess. Therefore, the Persson et al. (2003) hypothesis may hold for countries out of this sample, although in the appendix we also fail to replicate this hypothesis using the full sample of countries in a fixed effects analysis.

Second, it may not be surprising that there has been little change in the CCE scores of these
countries. None are very democratic, so that citizens may not be able to vote out legislators affiliated with a strong ruling party, and the locus of corruption may be the leaders of these parties rather than individual politicians. Togo did not hold an election under the new rules until 2013 because of a longstanding stalemate between the opposition and government. The situation in Kazakhstan is somewhat better in that although the ruling party held all legislative seats following the 2007 elections under the previous electoral system, some opposition parties were able to win seats in the highly disputed elections of 2012. Similarly, Kyrgyzstan adopted various constitutional reforms in 2007 following the Tulip Revolution, but the ruling party won an overwhelming majority of seats in the legislature. Changes in the electoral systems may have greater effects in more democratic countries.

Third, not enough time may have passed since the electoral system change for the effects to accrue in these ten countries, and this may be especially true for the plurality countries for which we only have a few post-treatment years. Our analysis may be replicated once more data becomes available on these countries. Fourth, we have adopted a nonparametric approach in order to reduce our sensitivity to measurement problems with CCE, and we find similar results with Transparency International’s CPI. However, these measures rely on surveys of experts for their perceptions of corruption, and more objective measures of corruption may reverse our findings.

Finally, synthetic control and difference-in-differences approaches may not have eliminated all bias due to unmeasured confounding, and the consequences are more problematic with worse measurement. We can conduct a Rosenbaum (2002) style sensitivity analysis on these results which would raise the reported p-values. However, such a sensitivity analysis would only weaken our confidence that PR might reduce corruption for these countries. It would not strengthen the evidence for the Persson et al. (2003) hypothesis.

\[^{10}\text{However, note that the data ends in 2011.}\]
6 Conclusion

Following the pioneering work of Persson, Tabellini and Trebbi (2003), scholars have proposed several theories linking electoral systems to levels of corruption. But how much empirical support these arguments have found has been debated, due to the difficulty of measuring corruption, non-random assignment of electoral systems, and the combination of different treatment effects.

This paper addressed each of these empirical challenges and found no support for the hypothesis that PR systems cause corruption. Rather, transitions to PR in our sample are weakly associated with lower levels of corruption. First, we addressed the issue of timing by focusing on individual country transitions from one system to another. Second, we more clearly defined the treatment effect by comparing treated countries only to countries with the same electoral system as the treated country had before its reform. Third, we address non-random treatment assignment by using synthetic control and differences-in-differences approaches. Finally, we partially addressed concerns about measurement of the outcome by using the nonparametric signed rank statistic.

The immediate implication is a call for a re-examination of the dominant view that plurality systems cause lower levels of corruption and that PR cause higher levels, the theory implicit in the South African petition for electoral reform. At a minimum, our analysis suggests that this is unlikely to be the case in new democracies and poorer countries, and we find some evidence that PR may even cause an increase in corruption. More generally, we should reconsider the conditions under which citizens hold elected leaders accountable for corruption, as this could yield stronger theory for the effect of electoral systems on corruption.
References


