# Ethnic winning coalitions and the political economy of aid*

Nils-Christian Bormann† Vera Z. Eichenauer‡
Simon Hug§

Department of Politics, Exeter University

KOF Swiss Economic Institute, ETH Zurich

Département de science politique et relations internationales, Université de Genève

Paper prepared for presentation at the Annual Meeting of
Midwest Political Science Association (Chicago, April 5-8, 2017)

**Abstract**

Selectorate theory provides an elegant and encompassing theoretical framework that makes predictions for several important puzzles in research on democracies and autocracies. Yet several critics lament shortcomings in the measurement of its key concepts: the size of the selectorate $S$ and the winning coalition $W$. We suggest an alternative that exploits information on the power status and population shares of ethnic groups around the world. Specifically, we identify the size of the selectorate as the sum of groups' population shares that do not suffer from political discrimination, and the size of the winning coalition as the cumulative population share of those ethnic groups represented in a state's executive. Our proposal improves on existing work by providing a continuous operationalization of $W$ and $S$, and thus to seamlessly bridge democratic and autocratic regimes, by not yielding observations in which the size of $W$ exceeds the size of $S$, and by ensuring that $S$ is strictly positive. We illustrate the usefulness by retesting and extending the claim that regimes with smaller winning coalitions receive higher levels of aid.

*An earlier version of this paper was prepared for the 10th Annual Conference on The Political Economy of International Organizations (Bern, January 12-14, 2016) and a seminar at the Hertie School of Governance. Helpful comments by the discussants Christina Schneider and Justin Vasalek and those by Patrick Bayer, Simone Dietrich, Iain Osgood, Peter Rosendorff, and Randall Stone as well as Alastair Smith's help with data issues are gratefully acknowledged.

† Department of Politics, Amory Building, Rennes Drive, Exeter, Devon, EX4 4RJ, United Kingdom; phone: +44 1392 723283; email: N.Bormann@exeter.ac.uk

‡ KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland Tel.: +41 - 44 - 632 28311 ; E-Mail: eichenauer@kof.ethz.ch

§ Département de science politique et relations internationales, Faculté des sciences de la société; Université de Genève; 40 Bd du Pont d'Arve; 1211 Genève 4; Switzerland; phone ++41 22 379 89 47; email: simon.hug@unige.ch

# 1 Introduction

Selectorate theory, as most fully articulated in "The logic of political survival" (Bueno de Mesquita, Smith, Siverson and Morrow, 2003), has created both new insights and solid theoretical underpinnings for such diverse research areas as the democratic peace and foreign aid allocation in international relations or autocratic leader survival and democratization in comparative politics. Its elegant setup relies on the relative importance of the population in charge of selecting the leader of a country, the selectorate $S$, and the population supporting the winner, namely the winning coalition $W$, and as a consequence permits parsimonious characterizations of political regimes. Thus, this theoretical approach has inspired a large set of research projects and contributions relying on the concepts proposed in this theory.

Bueno de Mesquita, Smith, Siverson and Morrow's (2003) theoretical approach and especially their empirical strategy have, however, attracted criticism. On a conceptual level some authors question whether in authoritarian regimes the notions of selectorate and winning coalitions even follow from institutional rules (e.g., Gallagher and Hanson, 2015). Clarke and Stone (2008) raise issues that undermine the empirical tests of selectorate theory such as the close link between the chosen indicators and a frequently used measure of democracy (for a more general critique of the indicators see Kennedy, 2009).

While even Bueno de Mesquita, Smith, Siverson and Morrow (2003, 133) concede the crudeness of their operationalization of the two key concepts, which draws on categorical information on political regimes (Banks, Day and Mueller, 1996; Marshall, Gurr, Davenport and Jaggers, 2002), no viable alternatives have been forthcoming so far. In this paper, we propose a possibly more accurate way to determine the size of the selectorate and the winning coalitions by drawing on the power-status of ethnic groups as recorded in the Ethnic Power Relations (EPR) dataset (Cederman, Wimmer and Min, 2010; Vogt, Bormann, Rüegger, Cederman, Hunziker and Girardin, 2015). As these measures directly refer to population shares, they can serve as basis for continuous measures of regime types, as suggested at the conceptual level by Bueno de Mesquita, Smith, Siverson and Morrow (2003, 72).

Admittedly, our exclusive focus on ethnic groups and their access to power presupposes that ethnicity is an important structuring factor in politics in general and in crucial policy areas in particular. Justifying this assumption below, we point to three key advantages of our alternative measures of $W$ and $S$: First, they seamlessly connect democracies and autocracies at both the theoretical and the empirical level and thus provide continuous operationalizations. Second, our measurements do not yield observations in which the size of the winning coalition exceeds the size of the selectorate, and third $S$ is always strictly positive.

We illustrate the usefulness of our indicators by replicating analyses that apply selectorate theory to allocation decisions of foreign aid (Bueno de Mesquita, Smith, Siverson and Morrow, 2003; Bueno de Mesquita and Smith, 2009), an area in which the selectorate theory has been applied with increasing frequency (see also Bueno de Mesquita and Smith, 2007; Bueno de Mesquita and Smith, 2016) and in which ethnicity plays also a significant role (see for instance Brown, Stewart and Langer, 2010; Briggs, 2012; Briggs, 2014; Jablonski, 2014; Dreher, Fuchs, Hodler, Parks, Raschky and Tierney, 2016). Our replications demonstrate that our proposed measures provide a valid alternative to existing measures. More specifically, we employ a variety of strategies including cross-validation to show that our measures provide a better fit to the data.

In the next section we briefly review the main building blocks of the selectorate theory and discuss both theoretical applications and empirical challenges. In section three, we propose alternative measures for both the selectorate $S$ and the winning coalition $W$ based on information on ethnic groups' access to political power. We then apply these measures to replicate a series of existing studies that assess how selectorate theory contributes to explaining foreign aid allocation decisions and do so also for a more up to date dataset. We conclude in section five by highlighting the advantages and limitations of our new conceptualization of the size of selectorates and winning coalitions.

## 2 The selectorate theory and its empirical challenges

Selectorate theory provides an elegant way to compare political regimes across the full spectrum from autocracies to full-fledged democracies. Thus, Gallagher and Hanson (2015, 368) argue that the "[selectorate] theory's most elegant innovation is to create a logic of accountability that links the policy outputs of rulers in all types of polities to the sizes of their winning coalitions..."

The fundamental assumption behind selectorate theory holds that national leaders aim to remain in power by distributing public and private goods to their supporters.

The mix between private and public goods depends on the relative size of the selectorate $S$ and the winning coalition $W$. Bueno de Mesquita, Smith, Siverson and Morrow (2003, 42) describe the selectorate as the subset of the population ". . . whose endowments include the qualities or characteristics institutionally required to choose the government's leadership and necessary for gaining access to private benefits doled out by the government's leadership." They define the winning coalition ". . . as a subset of the selectorate of sufficient size such that the subset's support endows the leadership with political power over the remainder of the selectorate as well as over the disenfranchised members of the society" Bueno de Mesquita, Smith, Siverson and Morrow (2003, 55).

Thus, selectorate theory predicts that leaders with small winning coalitions $W$ will focus on providing private goods to their limited number of supporters, while those faced with a large $W$ will provide more public goods. This basic logic is used by Bueno de Mesquita, Smith, Siverson and Morrow (2003) to explain a wide range of political outcomes, from peace and war (see also Bueno de Mesquita, Morrow, Siverson and Smith, 1999), to public goods provision and taxation (Bueno de Mesquita, Downs and Smith, 2017 (forthcoming)), leader survival (Bueno de Mesquita and Siverson, 1995), foreign aid (see below) etc. Gallagher and Hanson (2015, 368) quote from these authors' website that they consider it as a "power tool for explaining politics."

While the simplicity of the argument and the encompassing nature of its implications explain the attractiveness of this theory,[1] several scholars criticize the operationalization of the crucial elements of $W$ and $S$.[2] While Bueno de Mesquita, Smith, Siverson and Morrow (2003, 72) emphasize that ". . . $W$ and $S$ . . . are conceptually continuous variables . . ." (see also Bueno de Mesquita and Smith, 2007, 255), all their empirical applications effectively rely on ordered categorical variables (e.g. Bueno de Mesquita, Smith, Siverson and Morrow, 2003; Bueno de Mesquita and Smith, 2007, 2009$a$, 2009$b$, 2016 (forthcoming)), as they rely on components of the Polity IV democracy measure (Marshall, Gurr, Davenport and Jaggers, 2002) and the Banks dataset on political institutions (Banks, Day and Mueller, 1996). Specifically, Bueno de Mesquita and Smith (2009, 323) define their five-point measure of the winning coalition $W$ as:

"$W$ is normalized to vary between 0 and 1, with 1 representing the most

[1]According to Google Scholar almost 3,500 studies cite Bueno de Mesquita, Smith, Siverson and Morrow (2003) alone as of February 4th, 2017.

[2]Others critiques address the theory. For a valuable summary refer to Gallagher and Hanson (2015).

democratic countries and 0 the most autocratic. The estimate of winning coalition size relies on the Polity data [...] components REGTYPE (regime type), XRCOMP (the competitiveness of executive recruitment), XROPEN (the openness of executive recruitment), and PARCOMP (competitiveness of participation). One point is added to the index of $W$ for each of the following conditions: if the REGTYPE is nonmilitary, if XRCOMP is greater than or equal to 2 (meaning the chief executive is not chosen by heredity or in rigged, unopposed elections), if XROPEN is greater than 2, and if PARCOMP equals 5 (indicating the presence of a competitive party system)."

In their seminal book Bueno de Mesquita, Smith, Siverson and Morrow (2003, 134f) describe their measure of $S$ as follows:

> "We use a[nother] POLITY variable, Legislative Selection (LEGSELEC), as an initial indicator of $S$ . . . We divide LEGSELEC by its maximum value of 2 so that it varies between 0 and 1 . . ."

Table 1 summarizes these operationalizations of $S$ and $W$ which, even the authors concede, constitute crude ways to measure the central concepts and fail to produce continuous measures. Figure 1 depicts the relationship between $S$ and $W$ for the observations covered in Bueno de Mesquita, Smith, Siverson and Morrow (2003).[3] First, it illustrates the non-continuous distribution of the values for $S$ and $W$. Second, it reveals another problematic feature of the relationship between $S$ and $W$. By definition (see above), the winning coalition should always be a subset of the selectorate. Empirically, this means that all observations should fall below the diagonal depicted in Figure 1, which would imply that $W < S$. As the graph shows, however, this condition fails to hold in a considerable number of cases (1,513 out of 12,459). As long as we focus on $W$ or $S$ separately these cases constitute a lesser problem. However, many empirical applications in Bueno de Mesquita, Smith, Siverson and Morrow (2003) also rely on the so-called loyalty norm measured by $W/S$. Here the observations above the diagonal become more problematic because this fraction should have

[3]The data used is the replication dataset `bdm2s2_nation_year_data.dta` obtained at `http://www.nyu.edu/gsas/dept/politics/data/bdm2s2/Logic.htm` (accessed August 8, 2016). While this data provides values for $W$ from 1763 to 1999, only in 1815 the first value for $S$ appears in the data set. For this reason only data from the 1815-1999 period are used to produce Figure 1

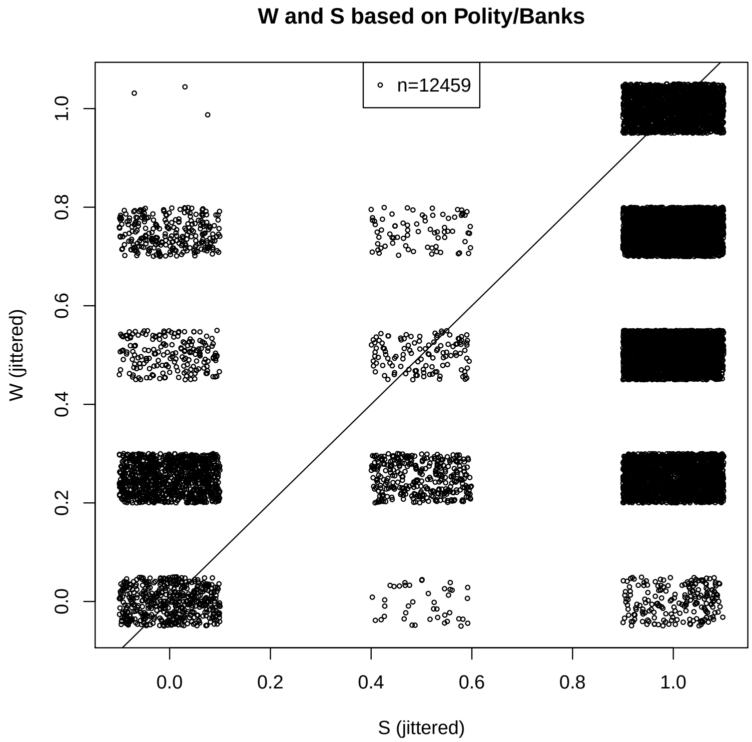


Figure 1: $W$ and $S$ based on POLITY (1815-1999, all POLITY countries)

an upper limit of 1.[4] Moreover, in some instances $S$ equals zero, and as a consequence Bueno de Mesquita, Smith, Siverson and Morrow (2003, 135) adjust the ratio $W/S$ on an ad-hoc basis. Gallagher and Hanson (2015, 376) understandably criticize the lack of theoretical justification for this procedure.[5]

[4]One might consider to rescale $W$ to tackle situations where it exceeds $S$. However, since sometimes $S = 0$, no rescaling method exists that would leave any $W > 0$.

[5]Among the observations depicted in Figure 1 the revised formula for $W/S$ generates for approximately 2,000 observations (out of 12,459) values larger than 1.

| Concept | Definition | Criteria | Bueno de Mesquita, Smith, Siverson & Morrow (2003) | | Ethnic Power Relations | |
|---|---|---|---|---|---|---|
| | | | **Measure**[i] | **Criteria ok?** | **Measure** | **Criteria ok?** |
| Selectorate $S$ | % of population *potentially* relevant for selecting the leader | Continuous variable; $S \in [0, 1]$ | $LEGSELEC/2$ | No, non-continuous variable | $S = 1-$ population share belonging to a discriminated ethnic group | Yes |
| Winning coalition $W$ | % of population selecting the leader | Continuous variable; $W \in [0, S]$ | One point is added to a cumulative index for each of the following conditions that is fulfilled: (i) REGTYPE is nonmilitary; (ii) XRCOMP $\geqslant 2$; (iii) XROPEN $\geqslant 2$; and (iv) PARCOMP $= 5$. The resulting sum is then divided by 4. | No, non-continuous variable; sometimes $W \notin [0, S]$ | $W$ = population share of ethnic groups included in executive | Yes |
| Loyalty norm $W/S$ | | Continuous variable; $W/S \in [0, 1]$ | | No, non-continuous variable; sometimes $W > S$; sometimes $W/S$ is undefined | $W/S$ | Yes |

*Notes*: LEGSELEC refers to legislative selection and comes from the Banks data (according to Bueno de Mesquita and Smith (2009) LEGSELEC is from the Polity IV dataset.); REGTYPE refers to regime type from Banks data (according to Bueno de Mesquita and Smith (2009) REGTYPE is from the Polity IV dataset); XRCOMP refers to the competitiveness of executive recruitment (Polity IV dataset); PARCOMP refers to the competitiveness of participation (Polity IV dataset).

Table 1: Overview over the measures

Beyond these inherent issues of measurement validity, several authors criticize the empirical tests in the work of Bueno de Mesquita, Smith, Siverson and Morrow (2003). Raising both conceptual and empirical concerns, Clarke and Stone (2008) criticize the way in which these authors deal with measures of the winning coalition and the Polity IV index of democracy. Since Bueno de Mesquita, Smith, Siverson and Morrow (2003) argue that both the size of the winning coalition $W$ and the degree of democracy in a country affect the provision of public goods, they wish to account for this latter confounding factor. Relying on components of the Polity IV scale as a source for the measures for $W$ and $S$ on the one hand and using all components for their measure of democracy on the other, proves to be problematic (Clarke and Stone, 2008). More specifically, Bueno de Mesquita, Smith, Siverson and Morrow (2003, 137) consider democracy to be endogenous to public good provision and therefore regress democracy on $W$ to use the residuals from this regression as a proxy for democracy. This amounts to ensuring that $W$ picks up all the shared effects of democracy on the dependent variable, as the residuals, by definition, are orthogonal to $W$. Clarke and Stone (2008) show, however, that this fix induces omitted variable bias. Correcting the problem yields much weaker support for implications derived from the selectorate theory.

Clarke and Stone (2008) offer a possible solution and propose an alternative measure for $W$ in democracies, namely the size of the governing coalition as measured by Powell (2000). Using this measure in conjunction with the Polity Democracy measure (Marshall, Gurr, Davenport and Jaggers, 2002), fails to produce support for the selectorate theory. Although Morrow, Bueno de Mesquita, Siverson and Smith (2008) acknowledge the econometric problems in their original analysis (Bueno de Mesquita, Smith, Siverson and Morrow, 2003, chapters 4 and 5), they hold that Clarke and Stone's (2008) solution is inadequate.[6] In response, Morrow, Bueno de Mesquita, Siverson and Smith (2008) pursue an alternative empirical strategy by operationalizing democracy only through the Polity IV component that is not used in the operationalization of $W$, namely the constraints on the executive.

While re-partitioning the Polity IV index into $W$ and a democracy residual solves the direct econometric concerns, it fails to address the deeper conceptual concerns. In fact, the new strategy simply repeats the earlier approach of rededicating a measure of

[6] Morrow, Bueno de Mesquita, Siverson and Smith (2008, 396) also reject Clarke and Stone's (2008) use of the size of the governing coalition as measure of $W$, presumably on the grounds that the latter is a behavioral measure and not an institutional one (information provided by Randall Stone in a personal communication, September 13, 2016). Note that our proposed measures also rely on behavioral aspects. As we discuss below, however, these behavioral aspects are strongly affected by institutional provisions.

democracy as an indicator for winning coalition size. Gallagher and Hanson (2015, 376, footnote 7) express this concern by stating that "[t]he more fundamental problem is that $W$ does not represent winning coalition size to begin with." There are two more straightforward ways to address the issue that regime type and the size of the winning coalitions both rely on Polity IV and influence public good provision: To propose alternative measures for $W$ and $S$ that do not rely on components of the Polity indicator (Gallagher and Hanson, 2015) or to replace the measure for democracy by an alternative, for instance the one presented by Alvarez, Cheibub, Limongi and Przeworski (1996) (see also Cheibub, Gandhi and Vreeland, 2010).

While Clarke and Stone's (2008) critique mostly targets the application of the selectorate theory in democracies, Gallagher and Hanson (2015) raise related concerns regarding the applicability of this theory in autocracies. Leaving aside further conceptual concerns, we will focus on the issues that deal with the use of Polity components to measure $W$ and $S$ in autocracies. Specifically, Bueno de Mesquita, Smith, Siverson and Morrow (2003) emphasize the institutional basis of their operationalization of $S$ and $W$ as a key advantage of their approach. However, Gallagher and Hanson (2015, 374) argue that in authoritarian regimes leadership changes usually depend less on formal institutions but follow more complex coalitional configurations. This critique reflects research that sees authoritarian institutions as an outcome rather than a determinant of elite coalition building and violent struggle (Slater, 2010; Pepinsky, 2014). Therefore, the measures for $W$ and $S$ proposed by Bueno de Mesquita, Smith, Siverson and Morrow (2003) have to be questioned in the context of authoritarian regimes, as "... authoritarian politics . . . often do not have any institutionalized structure for leadership selection or transition" (Gallagher and Hanson, 2015, 368). In China and Russia, for example, "the selectorate is not well defined even in these highly institutionalized authoritarian systems" (Gallagher and Hanson, 2015, 371).[7]

# 3 Winning coalitions, the selectorate and ethnic power relations

In response to the various criticisms leveled at selectorate theory, and especially the empirical operationalization of the key concepts, we propose alternative measures of

[7]More conceptual issues regarding the use of the selectorate theory to study authoritarian regimes are raised by Kennedy (2009) and Marcum and Brown (2016) (see also Gallagher and Hanson, 2013).

$S$ and $W$ that (i) span democratic and autocratic regimes and yield a continuous operationalization, (ii) do not produce observations in which the winning coalition size exceeds the size of the selectorate, and (iii) yield strictly positive values for the size of the selectorate. Our proposal draws on information on ethnic groups' leaders access to executive power and the population share of their groups. We are cognizant that basing measures of $S$ and $W$ on information on ethnic groups assumes that politics centers around ethnic divisions, and that ethnicity means the same across different contexts. We argue that ethnicity indeed constitutes a globally relevant cleavage in the time period that we study due to the ubiquity of the territorial state since the end of colonialism, and its close link to nationalism exported by Western states to their former colonies. Nationalism is most famously defined by Gellner (1983, 1) as the ideology that holds that "the political and national unit should be congruent." Nationalism thus often divides individuals living in the same state along ethnic lines. While different definitions of nationhood exist, the most common one is based on membership in ethnic groups (Mann, 2005). Hechter (2000, 62) explains the reason for the special role of ethnicity: "[C]ultural uniformity helps to facilitate, and to legitimize, direct rule."

For similar reasons ethnicity now features prominently in research that shares many assumptions with selectorate theory. Easterly and Levine's (1997) influential study considered ethnic fragmentation as one reason for the limited growth in African economies. Underlying these arguments is on the one hand the idea that interactions across ethnic boundaries are more difficult and on the other that fragmentation also leads to differential access to power. Specifically, the visibility of ethnic markers enables elites to form coalitions that exclude members of some ethnic groups while giving preferential treatment to their own (Bates, 1974; van der Veen and Laitin, 2012).[8] Although Kasara (2007) argues that national leaders find it easier to tax their co-ethnics due to informational advantages, most scholars argue that ethnic groups benefit when their elites hold government power (Burgess, Jedwab, Miguel, Morjaria and Padro i Miquel, 2015; Morelli and Rohner, 2015; Alesina, Michalopoulos and Papaioannou, 2016). A recent study by Weidmann, Benitez-Baleato, Hunziker, Glatz and Dimitropoulos (2016) illustrates the interplay of political inequalities and the provision of public goods by showing that internet traffic is much lower in geographic areas popu-

[8] Other scholars explore the consequences of these ethnic or horizontal inequalities, particularly on civil war (Gurr, 2000; Wimmer, 2002; Stewart, 2008; Cederman, Wimmer and Min, 2010; Cederman, Gleditsch and Buhaug, 2013).

lated by ethnic groups excluded from power (see also Baldwin and Huber, 2010).[9]

Relatedly, several recent studies show that coalitions among ethnic groups in power follow a logic of survival akin to Bueno de Mesquita, Smith, Siverson and Morrow's (2003) argument in both democracies and autocracies (for a detailed study, see Bormann, 2014). Focusing on dictatorships, Beiser and Metternich (2016) show that the coalitions of ethnic groups forming in autocracies balance off the threat of a coup from the inside and challenges from the outside (for similar and related arguments, see Roessler, 2011; Roessler, 2016; Roessler and Ohls, 2017 (forthcoming)). Roessler (2016), drawing on a rich literature on African states, even argues that in the weak states of this region (main recipients of aid) the alliances amongst leaders of ethnic groups (and not institutions) are central to understand politics, coups and civil wars. Thus, we argue that information on access to executive power enjoyed by ethnic groups might offer better measures that seamlessly bridge autocracies and democracies.

Many of the studies discussed above rely on information from the EPR dataset (Vogt, Bormann, Rüegger, Cederman, Hunziker and Girardin, 2015).[10] This dataset codes politically relevant ethnic groups between 1946 and 2013 in all states with a population in excess of 500'000. The EPR data consider linguistic, religious, racial, and caste differences, but not clan cleavages, for all groups, whose leaders advance political claims on their behalf in the national arena, or for groups discriminated politically by the state. In addition, the data provide information on groups' population share and their leaders' access to executive power as shown in Table 2.[11]

Whereas the monopoly and dominant categories characterize monoethnic governments, senior and junior partners mark power-sharing regimes. To qualify for an inclusion coding, groups need to have effective and not just token representation in the high-

[9] Other researchers report similar findings with regard to electricity (e.g., Baskaran, Uppal and Min, 2015; Min, 2015; Kroth, Larcinese and Wehner, 2016).

[10] EPR, Luc Girardin, Philipp Hunziker, Lars-Erik Cederman, Nils-Christian Bormann, and Manuel Vogt. 2015. $GROW^{up}$ - Geographical Research On War, Unified Platform. ETH Zurich. `http://growup.ethz.ch/` (accessed August 8, 2016 Vogt, Bormann, Rüegger, Cederman, Hunziker and Girardin, 2015).

[11] Obviously, there are other sources on ethnic groups, like for instance the "Atlas Narodov Mira" (Bruk, 1964) used by Easterly and Levine (1997) to determine the fragmentation of African societies. This source, however, as many other lists of ethnic groups, fails to give any information on the latter's access to power, which is key for measures of $S$ and $W$. Relatedly, the data on "Minorities at Risk" (Marshall, Gurr, Davenport and Jaggers, 2002) offer detailed information on discriminated and/or mobilized groups, but fail to provide information on other groups. An extension of this dataset, namely A-MAR (All-MAR, see Birnir, Wilkenfeld, Fearon, Laitin, Gurr, Brancati, Saideman, Pate and Hultquist, 2015) proposes to cover all "socially relevant groups" but offers only detailed information for a sample of these groups.

|  | Power Status | Group-Years | Share |
|---|---|---|---|
| Included | Monopoly | 1,944 | 0.053 |
|  | Dominant | 2,303 | 0.063 |
|  | Senior Partner | 3,522 | 0.096 |
|  | Junior Partner | 6,997 | 0.190 |
| Excluded | Powerless | 15,634 | 0.425 |
|  | Self-excluded | 592 | 0.016 |
|  | Discriminated | 5,822 | 0.158 |

Table 2: Categories and distribution of power access in the EPR dataset, 1946–2013

est national executive body. The latter comprises communist central committees, military juntas, or royal courts in dictatorships, and presidential and parliamentary cabinets in democracies. Groups excluded from the national executive fall into one of three categories: powerless groups, which lack influence in the executive, self-excluded groups that declared independence from the central government, and discriminated groups, which face "active, intentional and targeted political discrimination by the state" such as the denial of voting or even citizenship rights (Vogt, Bormann, Rüegger, Cederman, Hunziker and Girardin, 2015, 1331). Combining information on groups' power status and their relative size, we propose novel measures for $S$ and $W$ along the following lines:[12]

$$\begin{aligned} S &= 1 - \text{population share belonging to (a) discriminated ethnic group(s)} \\ W &= \text{population share of ethnic groups included in executive} \end{aligned}$$

The explicit focus of the EPR-data on executive power and the link of political power to distributional outcomes justifies our measure of the winning coalition $W$.[13] Operationalizing the selectorate proves a greater challenge. EPR codes discriminated groups exactly because the state actively excludes them from this power. Therefore members of discriminated groups should not be included in the selectorate. Powerless groups, on the other hand, are defined as "simply not represented in the execu-

[12]The EPR-dataset does not provide information for twenty-four countries without politically relevant ethnic differences such as Sweden or Denmark. In these cases, we assume that the country is homogeneous and $S$ and $W$ equal one. In our empirical analyses we systematically assess whether dropping these observations changes our insights.

[13]Heger and Salehyan (2007), for instance, assume that for study on Africa the winning coalition $W$ corresponds to the population size of the head's of state ethnic group. Francois, Rainer and Trebbi (2015), on the other hand, consider the ethnic background of all cabinet ministers in Africa in their study of power-sharing.

tive" (Vogt, Bormann, Rüegger, Cederman, Hunziker and Girardin, 2015, 1331). As such they may theoretically influence the composition of the winning coalition. Self-excluded groups fall somewhere in-between these two categories but, as Table 2 indicates, they are so rare that assigning them to the selectorate hardly matters. In the following analyses we do not count them towards the selectorate. By proceeding in this way our measures of $S$ and $W$ are continuous and span the whole spectrum from autocracies to democracies per our first criterion.[14]

To illustrate our measure we discuss a few select cases and compare our measure with the original ones proposed by Bueno de Mesquita, Smith, Siverson and Morrow (2003). A country frequently in the news over the last years because of its exclusionary government composition is Syria. Our proposed measure of $S$ and $W$ for 1999 (this is the last year covered by the replication data for Bueno de Mesquita, Smith, Siverson and Morrow, 2003) is 0.92, resp. 0.13.[15] Our coding of $S$ derives from the discrimination of the Kurds with a population share of 0.08 (see `https://growup.ethz.ch/pfe/Syria`). Since EPR considers President Assad's government to consist exclusively of his fellow Alawis, $W$ takes on the respective population share of 0.13. Bueno de Mesquita, Smith, Siverson and Morrow's (2003) values for $S$ and $W$ in 1999 are, on the other hand, 1 resp. 0.5. While these values show similar tendencies they diverge nevertheless quite considerably from ours.

A case where there is no divergence is Switzerland. Both Bueno de Mesquita, Smith, Siverson and Morrow's (2003) values as well as ours are equal to 1. For our measures, this comes about by the fact that the EPR data consider German-, French- and Italian-speakers as ethnic groups included in government, while the remaining native and the high-share of non-naturalized population is considered irrelevant (see `https://growup.ethz.ch/pfe/Switzerland`).

A slightly more systematic comparison can be achieved by looking at cases where the difference in the two sets of measures is maximal. It turns out that in Bueno de Mesquita, Smith, Siverson and Morrow's (2003) replication set this occurs for a few cases where $S$, resp. $W$ is equal to 0, while our measures for $S$, resp. $W$ indicate a value of 1. Table 3 lists these cases and shows that with the exceptions of Argentina

[14]The correlations between Polity IV on one hand and $S$ and $S_{EPR}$ on the other are .35 and .33 respectively. They thus do not give any cause for concern in terms of collinearity as they might do with Bueno de Mesquita, Smith, Siverson and Morrow's (2003) measure of $W$ which correlates strongly with the Polity IV scale (.82).

[15]EPR's coding of Syria remained the same until 2011, when the Syrian civil war started (see `http://www.ucdp.uu.se/#country/652`).

Fiji, Haiti, Nicaragua, Oman and South Korea only African countries appear in this list.

| Country | Year | S | S_epr | W | W_epr |
|---|---|---|---|---|---|
| Algeria | 1994-1996 | 0.00 | 1.00 | | |
| Argentina | 1970 | 0.00 | 1.00 | | |
| Burkina Faso | 1991 | 0.00 | 1.00 | 0.00 | 1.00 |
| Burundi | 1991-1992 | 0.00 | 1.00 | 0.00 | 1.00 |
| | 1998- 1999 | 0.00 | 1.00 | | |
| Chad | 1991 | 0.00 | 1.00 | | |
| Congo | 1997- 1998 | 0.00 | 1.00 | | |
| Fiji | 1991 | 0.00 | 1.00 | | |
| Ghana | 1972-1976, 1981-1985, 1987, 1989-1991 | 0.00 | 1.00 | 0.00 | 1.00 |
| Haiti | 1987 | 0.00 | 1.00 | | |
| Lesotho | 1987-1990 | 0.00 | 1.00 | 0.00 | 1.00 |
| Madagascar | 1972 | 0.00 | 1.00 | | |
| Mali | 1980 | 0.00 | 1.00 | | |
| Morocco | 1972- 1975 | 0.00 | 1.00 | | |
| Nicaragua | 1971 | 0.00 | 1.00 | | |
| Nigeria | 1978, 1984- 1987 | 0.00 | 1.00 | | |
| Oman | 1971- 1980 | 0.00 | 1.00 | | |
| Sierra Leone | 1992 | 0.00 | 1.00 | | |
| | 1993-1995 | 0.00 | 1.00 | 0.00 | 1.00 |
| South Korea | 1972 | 0.00 | 1.00 | | |
| Swaziland | 1975 | 0.00 | 1.00 | | |

Table 3: Cases with maximal discrepancies in the values of $S$ and $W$

Table 3 already offers some illustration for cases in which our values for $S$ or $W$ are maximal, namely 1. For this reason it is also helpful to highlight cases in which the values for these two variables are minimal. We find the lowest values both for $S$ and $W$ in Liberia, namely 0.02, in the whole post World War II period until 1980. These values reflect the EPR coding of Americo-Liberians as holding the monopoly of power and discriminating against all other indigenous ethnic groups in Liberia which were not granted voting rights until 1985 (see `https://growup.ethz.ch/pfe/Liberia`). Thus, our measures consider the Americo-Liberians to constitute both the selectorate $S$ as well as the winning coalition $W$. This contrasts with Bueno de Mesquita, Smith, Siverson and Morrow's (2003) measures which for the same period are respectively 1 ($S$) and 0.5 ($W$), with the exception of 1980, where these values both drop to 0.

We find a similarly low value for our $W$ for Guinea-Bissau from 1974 until 1980. For these early years after independence EPR considers the Cape Verdeans, the ethnic group installed as the administrative elite by the former colonial power Portugal, as dominant and all other groups as powerless. Unlike in the Liberian case the excluded groups such as the Balanta constituted a large part of the army. Consequently, $S$ is equal to 1 for this country in this period, while $W$ equals also 0.02. Again, Bueno de Mesquita, Smith, Siverson and Morrow's (2003) measure differ quite dramatically as again for the whole period $S$ is equal to 1 and $W$ equal to 0.5, except in 1980 when $S$ drops to 0 and $W$ to 0.25.

To proceed more systematically we combine the EPR data and the data from Bueno de Mesquita and Smith (2009) and focus on the observations included in both datasets. For the time period 1960-1999 and the countries with a population in excess of 500'000, Figure 2 thus reproduces Figure 1 for the proposed measures $S$ and $W$ as well their original measures.[16] Excluding cases not captured in the EPR data does not fundamentally alter the distribution of $W$ and $S$ and there remain more than five percent of all observations that fall above the diagonal. In contrast, the EPR-based measures of $W$ and $S$ in the second panel satisfy the condition that $W < S$ by construction.

Comparing the two graphs suggests that the two sets of measures are related.[17] In most country-years, the selectorate includes the full population ($S = 1$), most of which is also part of the winning coalition (high values of $W$). Yet the EPR measures count far fewer cases with a tiny selectorate and a small winning coalition. It is possible that the EPR-based measures overestimates the winning coalition and selectorate size or that the original measures underestimated them (see the measure based on heads of state and the discussion on this in Heger and Salehyan, 2007).

# 4 Ethnic winning coalitions and aid

While we have argued that our measures of the selectorate and the winning coalition provide a more convincing conceptual fit than the ones of Bueno de Mesquita, Smith, Siverson and Morrow (2003), we recognize that we cannot prove the superiority of our measures empirically. In what follows we employ our measures in replications of analyses focusing on foreign aid allocation and provide at least suggestive evidence that our measures are preferable. This area of research has recently seen a series of applica-

[16]Figure 1 covers much more cases, as Bueno de Mesquita, Smith, Siverson and Morrow's (2003) replication data covers all polities starting at the end of the 18th century, while for Figure 2 we focus on the period 1960-1999, which is the basis for our first replication below.

[17]The correlations, while significant, are not particularly strong with 0.087 for $S$ and 0.174 for $W$. We also note that in the left-hand panel 843 observations out of 5105 appear above the diagonal, thus implying that that $S > W$

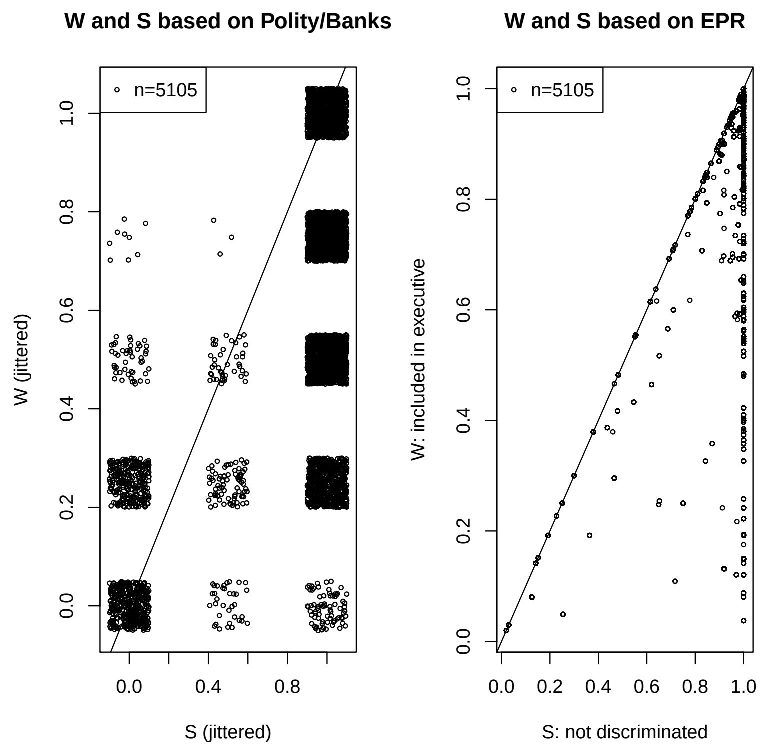


Figure 2: $W$ and $S$ based on Polity and EPR (1960-1999, population 500’000)

tions of the selectorate theory (Bueno de Mesquita, Smith, Siverson and Morrow, 2003; Bueno de Mesquita and Smith, 2007; Bueno de Mesquita and Smith, 2009; Bueno de Mesquita and Smith, 2016). In addition, several recent studies in this area also emphasize the role ethnicity (for instance of the head of state of the recipient) play in aid allocations (see for instance Brown, Stewart and Langer, 2010; Dreher, Fuchs, Hodler, Parks, Raschky and Tierney, 2016; Briggs, 2012; Briggs, 2014; Jablonski, 2014). Thus, we use two studies (Bueno de Mesquita, Smith, Siverson and Morrow, 2003; Bueno de Mesquita and Smith, 2009) and show how replacing the original measures for $S$ and $W$ with our proposed measures affect the existing results. In addition, by relying on more recent data a more conventional empirical specification we assess the effect of $W$ on bilateral aid allocation in a broader context. For all studies, by analyzing the fit of the empirical models and the residuals, as well as carrying out cross-validations, we hope to demonstrate that our measures are preferable to the original ones.

While the first study that we replicate only offers a brief analysis of aid recipients

(Bueno de Mesquita, Smith, Siverson and Morrow, 2003), more recent work by Bueno de Mesquita and Smith (2009) extends selectorate theory into the realm of foreign aid. The main theoretical reasons for donors to provide aid is to "buy" policy concessions from recipients. Recipients are more likely to make policy concessions if the additional resources provided by donors enhances their political survival if adjusting their policy position towards the donor's does not hurt this goal. We will assess the quality of our suggested operationalizations of $S$ and $W$ based on the original specifications, although the literature on aid allocation has converged on slightly different covariates. Moreover, the fit between the argument in Bueno de Mesquita, Smith, Siverson and Morrow (2003, 449) does not correspond very closely to the empirical implementation. Implicitly, Bueno de Mesquita and Smith (2009) acknowledge this by providing more developed arguments about the political economy of aid and by estimating substantially different specifications (see Table 6).

According to Bueno de Mesquita, Smith, Siverson and Morrow (2003), policy concessions are cheaper to buy from leaders with small $W$s so that the countries governed by such leaders are both more likely to receive any and more aid. In these regimes, leaders are especially eager for additional funds to cater private benefits to their winning coalition to enhance their prospects of political survival. According to our reading of Bueno de Mesquita, Smith, Siverson and Morrow (2003, 279), the so-called loyalty norm ($W/S$) should relate positively to the probability of receiving aid and aid receipts.[18] Finally, the authors hypothesize that the effective $S$ ($eS = S \times (1 - W)$ (Smith and Vreeland, 2010)), which takes high values in "rigged election autocracies," should also lead to disproportionate high amounts of aid. To empirically assess these hypotheses, Bueno de Mesquita, Smith, Siverson and Morrow (2003, 480) propose an analysis of the recipients of US bilateral aid. More specifically they estimate a Heckman selection model in which the selection equation estimates the likelihood that states receive aid from the US, while the outcome equation estimates the amount of aid that states receive. Table 4 reports our results.[19] Models 1-3 use the original proxies of $W$ and $S$, whereas Models 4-6 employ our EPR-based variables. Model 1 restates the results reported in Bueno de Mesquita, Smith, Siverson and Morrow (2003). Model 2 reports our replication of the same model based on the available replication dataset,

[18]In the original specification in column 1 of Table 4, the loyalty norm is included in the second stage only while $S$ is added to the first stage although Bueno de Mesquita, Smith, Siverson and Morrow (2003, 279) make no argument about its relationship with aid receipts.

[19]We thank Alastair Smith for his help in reproducing these results.

and Model 3 estimates that model only on countries with a population greater than 500,000 for which both the original proxies and the EPR measures exist.[20] Model 4 uses all cases from the EPR data whereas Model 5 uses the sample from Model 3. Finally, the sample in Model 6 only drops those twenty-four (donor and recipient) states which EPR characterizes as featuring no politically relevant ethnic divisions (e.g. Denmark and Sweden).

Although some differences exist, the results for Models 1 and 2 are very similar and yield a negative estimated effect of $W$ on the probability of a recipient receiving aid, which aligns with the theoretical expectations. Similarly, the coefficient for the recipient's $W$ exhibits the expected negative effect on the amount of aid received. Finally, the coefficient for the loyalty measure ($W/S$) is positive as expected. Model 3 relies on the shared sample and the results remain quite similar. Only the coefficient for $W$ in the selection equation becomes much smaller and fails to reach statistical significance.

Turning to the EPR-based measures of $S$ and $W$, the results in Model 4 reveal both similarities and differences. Whereas the results for $S$ in the selection equation and $W$ in the outcome equation remain in line with the predictions of selectorate theory, the coefficients for $W$ in the selection equation and for $S$ in the outcome equation reverse their sign, even if they barely miss conventional measures of statistical significance. According to these results, states with larger winning coalitions are more likely to be the recipients of US aid whereas states with larger selectorates receive higher amounts from this donor. Models 5 and 6 generally recover those effects. In Model 5 the positive effect of $W$ on the likelihood of receiving aid turns statistically significant but drops below significance again in states with politically relevant ethnic divisions (Model 6). In the outcome equation the coefficient for the effective $eS$ fails to reach statistical significance in Model 5 but passes it in Model 6.[21]

To evaluate the more substantial differences between the different proxies for $S$ and $W$, we compare Models 3 and 5 that rely on an identical sample. Therefore, the comparison of the log-likelihood gives us some indication of the relative fit of the two

[20]Of the 196 observations we lose by doing so all but fourteen stem from countries with population sizes smaller than half a million. The fourteen other cases correspond to nine from Germany, four from Namibia and one from the Comoros Islands.

[21]The subtle differences in estimation results between different sample specifications, particularly in Models 4-6, might derive from violations of the identification restriction of the Heckman-model, which requires that $S$ is sufficiently different from the effective $S$, as the former is excluded from the outcome equation.

models. As the two values suggest, using our measures for $S$ and $W$ leads to a higher value of the log-likelihood, which suggests a better fit.[22]

| | Original | Replication | | Replication EPR | | |
|---|---|---|---|---|---|---|
| | | | Adjusted n | All countries | Adjusted n | Relevant countries |
| | $Model$ 1 | $Model$ 2 | $Model$ 3 | $Model$ 4 | $Model$ 5 | $Model$ 6 |
| | | | Selection | | | |
| Recipient coalition $W$ | $-1.087^{*}$ | $-1.214^{*}$ | $-0.420$ | $0.961$ | $1.524^{*}$ | $0.654$ |
| | $(0.422)$ | $(0.406)$ | $(0.415)$ | $(0.555)$ | $(0.565)$ | $(0.628)$ |
| Recipient selectorate $S$ | $-1.533^{*}$ | $-1.274^{*}$ | $-2.140^{*}$ | $-3.319^{*}$ | $-3.933^{*}$ | $-2.710^{*}$ |
| | $(0.564)$ | $(0.511)$ | $(0.559)$ | $(0.718)$ | $(0.731)$ | $(0.770)$ |
| Ln(GDP) | $-2.188^{*}$ | $-1.967^{*}$ | $-2.036^{*}$ | $-2.084^{*}$ | $-2.116^{*}$ | $-2.040^{*}$ |
| | $(0.149)$ | $(0.124)$ | $(0.132)$ | $(0.125)$ | $(0.126)$ | $(0.129)$ |
| Government debt | $0.016^{*}$ | $0.016^{*}$ | $0.019^{*}$ | $0.015^{*}$ | $0.014^{*}$ | $0.014^{*}$ |
| | $(0.002)$ | $(0.003)$ | $(0.002)$ | $(0.002)$ | $(0.002)$ | $(0.002)$ |
| (Intercept) | $21.948^{*}$ | $19.706^{*}$ | $20.216^{*}$ | $20.822^{*}$ | $21.150^{*}$ | $20.146^{*}$ |
| | $(1.419)$ | $(1.200)$ | $(1.268)$ | $(1.165)$ | $(1.182)$ | $(1.210)$ |
| | | | Outcome | | | |
| Recipient coalition $W$ | $-269.082^{*}$ | $-228.419^{*}$ | $-353.663^{*}$ | $-166..983^{*}$ | $-166.235^{*}$ | $-136.291^{*}$ |
| | $(90.604)$ | $(124.000)$ | $(18.211)$ | $(29.050)$ | $(29.210)$ | $(31.391)$ |
| Effective $eS$ | $-38.071^{*}$ | $-46.321^{*}$ | $-42.237^{*}$ | $108.831$ | $136.583$ | $166.652^{*}$ |
| | $(11.023)$ | $15.080)$ | $(13.250)$ | $(77.851)$ | $(78.285)$ | $(85.174)$ |
| Ln(GDP) | $15.804^{*}$ | $26.537^{*}$ | $27.147^{*}$ | $25.333^{*}$ | $25.444^{*}$ | $26.343^{*}$ |
| | $(1.853)$ | $(2.483)$ | $(2.308)$ | $(2.164)$ | $(2.220)$ | $(2.458)$ |
| Government debt | $0.382^{*}$ | $0.664^{*}$ | $0.695^{*}$ | $0.602^{*}$ | $0.602^{*}$ | $0.678^{*}$ |
| | $(0.039)$ | $(0.054)$ | $(0.048)$ | $(0.047)$ | $(0.047)$ | $(0.054)$ |
| $W/S$ | $243.808^{*}$ | $210.417$ | $319.873^{*}$ | $261.473^{*}$ | $260.490$ | $291.689$ |
| | $(94.773)$ | $(129.696)$ | $(117.588)$ | $(54.231)$ | $(54.597)$ | $(59.976)$ |
| (Intercept) | $-60.965^{*}$ | $-138.563^{*}$ | $-146.131^{*}$ | $-268.556^{*}$ | $-278.916^{*}$ | $-339.304^{*}$ |
| | $(14.843)$ | $(20.076)$ | $(18.210)$ | $(78.470)$ | $(78.905)$ | $(86.142)$ |
| $N$ censored | 420 | 424 | 360 | 360 | 360 | 305 |
| $N$ uncensored | 1225 | 1222 | 1089 | 1105 | 1089 | 903 |
| $N$ | 1645 | 1646 | 1449 | 1465 | 1449 | 1208 |
| $\ell$ | $-7109.28$ | $-7491.148$ | $-6520.318$ | $-6567.836$ | $-6469.101$ | $-5398.329$ |

*Notes:* Standard errors in parentheses. $^{*}$ indicates significance at $p < 0.05$.
Censored observations are those country-years that do not receive any bilateral US aid. Models 2 and 3 replicate Bueno de Mesquita, Smith, Siverson and Morrow (2003). The dependent variable of the selection equation is one in country-years with positive US aid flows and zero otherwise. The dependent variable of the outcome equation is logged bilateral US aid. The unit of analysis is the recipient-year level.

Table 4: Heckman model of aid

Our analysis of model fit suggests that our measures improve the fit of the model to the data relative to the original measures of $S$ and $W$. In the following, we probe whether our critique of the original measures, focusing on implausible values, is to

[22]As the degrees of freedom are identical across the two models information criteria (which we do not report here) would lead to the same conclusion.

blame for the inferior performance of the original proxies. As discussed above, we argue that their lack of correspondence between the empirical and the theoretically expected values constitutes one of the main problems of the original measures. The problem becomes glaring when these measures suggest that the winning coalition $W$ is larger than the selectorate $S$. In such situations the loyalty norm $W/S$ would take values larger than one, which incidentally is also undefined for values of $S$ equal to zero. For this reason Bueno de Mesquita, Smith, Siverson and Morrow (2003) propose an alternative formula that leads to a defined value for the loyalty norm even if $S$ is equal to zero. Nevertheless, we would suspect that especially cases in which $W$ is larger than $S$ (or either of these two variables is equal to 0) are those that might be most problematic.

To assess this claim Table 5 reports the results of simple regressions using as dependent variable the absolute values of the residuals from the outcome equation of Model 3 from Table 4. The first model uses a dichotomous indicator for all cases in which $W$ exceeds $S$ for the original measures. The observations for which this is the case have significantly larger absolute values for the residuals than those for which this condition does not hold. In the second model we add two indicators for cases in which either $S$ or $W$ is equal to zero. In this model we find that the residuals for observations where $W$ is zero are on average smaller, while for the two other dichotomous indicators we find positive effects. For the cases where $S$ is smaller than $W$ we still find a positive and significant effect, suggesting that these cases are actually problematic both from a measurement perspective and a substantive one.

In the next two columns we report two equivalent models for the EPR proxies from Model 5 in Table 4. We expect that the problematic values of the original measures for $S$ and $W$ should exert a smaller or negligible effect on the absolute size of the residuals. This, however, is not completely the case. We find for these two models quite similar patterns, namely that having $W$ smaller than $S$ yields higher absolute values for the residuals, and that when taking into account whether $S$ and/or $W$ is zero, the latter has a negative effect on the absolute value of the residuals. While for the condition $S < W$ we actually find the expected pattern, namely smaller absolute residuals if our proposed measures are used, for the other two variables the effects are different. In cases where the original measures for $W$ yield the value of 0 using our measures results in even smaller residuals, while for cases where $S$ is 0, using our measure results in even higher absolute values of the residuals. In addition, Table 5

| | Replication | | Replication EPR | |
|---|---|---|---|---|
| | $Model\ 1$ | $Model\ 2$ | $Model\ 3$ | $Model\ 4$ |
| $S < W$ | $63.10^{*}$ | $51.84^{*}$ | $51.96^{*}$ | $33.61^{*}$ |
| | $(7.40)$ | $(17.91)$ | $(6.96)$ | $(16.82)$ |
| $W = 0$ | | $-22.39$ | | $-30.77^{*}$ |
| | | $(13.65)$ | | $(12.82)$ |
| $S = 0$ | | $10.19$ | | $17.03$ |
| | | $(16.30)$ | | $(15.31)$ |
| (Intercept) | $44.61^{*}$ | $45.68^{*}$ | $44.57^{*}$ | $45.90^{*}$ |
| | $(1.99)$ | $(2.06)$ | $(1.88)$ | $(1.94)$ |
| $N$ | 1089 | 1089 | 1089 | 1089 |
| Resid. sd | 63.37 | 63.30 | 59.61 | 59.45 |

*Notes:* Standard errors in parentheses.
$^{*}$ indicates significance at $p < 0.05$
The dependent variables of columns 1 and 2 and of columns 3 and 4 are the absolute value of the residuals from the outcome equation in columns 3 and 5 of Table 4 respectively.

Table 5: Residuals and values of $S$ and $W$

shows that the residuals for the models with our measures are substantially smaller on average (see estimates for the intercepts). This underlines the finding based on a comparison of the log-likelihoods from Table 4.

Another way to assess the performance of our new measures is to rely on cross-validation (Geisser, 1975; Efron, 1983; Arlot and Celisse, 2010, for a comprehensive survey). As the sample sizes for the analyses reported in Table 4 are relatively small, the choice between a $k$-fold and a Monte-Carlo cross-validation presents important trade-offs. While splitting the sample in $k$ folds and estimating the model $k$ times while omitting one of the folds that is then used for the calculation of the prediction error, is known to lead to unbiased estimates, it comes with the disadvantage of yielding high variances. Monte-Carlo cross-validation, i.e., randomly selecting the training set and calculating prediction errors on the remaining observations, may lead to biased estimates, as several observations might be used repeatedly in the estimations, but yields generally lower variances. For this reason, we carry out both $k$-fold and Monte-Carlo cross-validations with a broad set of parameters.[23] For each estimation round we

[23] More specifically, we carried out four Monte-Carlo cross-validations in which we drew either 500 or 1000 random samples of 10 or 50% of the original sample that we removed to form the test sample (resp. mc 500, 10; mc 1000, 10; mc 500, 50; mc 1000, 50). Then we also randomly subdivided the sample in 10, 50, 100, 500 or 1000 folds and estimated the model while removing one of the folds (10-fold, 50-fold, 100-fold, 500-fold and 1000-fold). Given the small sample size and the sensitivity of the Heckman selection model we also encountered some convergence problems for some of the random

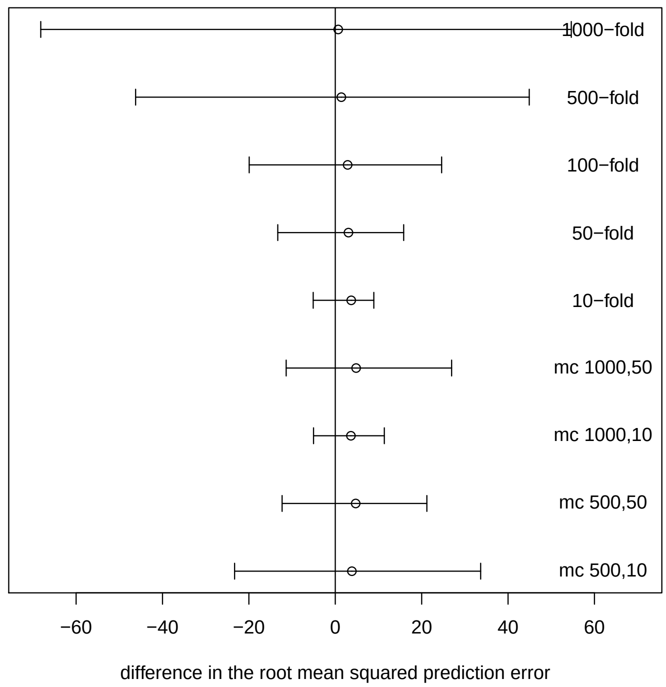


Figure 3: Comparison of prediction errors based on $k$-fold and Monte-Carlo cross-validations

estimated two models, one with the original measures for $S$ and $W$ and one using our own measures (and this on the same set of observations as illustrated in Table 4). We then used the estimates for each model to calculate predictions which we compared with the actual values of the dependent variable in the set of observations not used for estimation. Based on this we calculated the mean squared prediction error for both models and took the difference.

Figure 3 depicts our results. In the nine cross-validations the difference in the root of the mean squared prediction error between the estimates based on the original measures of $S$ and $W$ and those obtained based on our measures always yields a positive difference. Thus, on average, the prediction error is always larger when using the original measures for $S$ and $W$ than those we obtain based on our proposed measures. The 95% confidence interval, however, for all cross-validations comprises the value

samples in the Monte Carlo cross-validations. As these were only a handful, we removed them (without replacing them) from the calculation of the prediction errors.

of 0. Nevertheless, both the mean differences and the analyses based on the residuals (Table 5) suggest that our measures lead to improved predictions.

In a more recent article Bueno de Mesquita and Smith (2009) develop their argument regarding the "political economy of aid" (for reviews of this literature, see Wright and Winters, 2010; Fuchs, Dreher and Nunnenkamp, 2014) and provide more detailed analyses of bilateral aid flows for all DAC donors at the dyadic donor-recipient level rather than the recipient level and for the United States only. Drawing on their argument that foreign aid is used to extract policy concessions from recipient countries (Bueno de Mesquita and Smith, 2007), their theoretical model implies that rich countries with large $W$ will give more aid, and that increasing recipient coalition size $W$ initially leads to more aid, the relationship subsequently reverses until no more aid is given (Bueno de Mesquita and Smith, 2009, 321).[24] In this more recent analysis, the size of the effective selectorate $eS$ and the loyalty norm $W/S$ enter neither the argument about the likelihood of "selling" policy concessions nor the regression specification.

The empirical results for $W$ reported in this article are mixed at best (table 6, model 1). The $W$ of the donors has a significantly negative effect on (logged) aid amounts, which the authors attribute to the fact that in their OECD sample almost all countries are democracies and thus have a $W$ of 1. Regarding the effect for the $W$ of the recipients the relationship appears to be curvilinear but not in the direction predicted by the theory. More precisely for small values of the recipients' $W$ the effect on aid received is negative. As $W$ becomes larger, the amount of aid delivered increases at an accelerating rate. When replicating this model with the data made available by the authors we obtain largely the same results (table 6, model 2), and we depict the implied curvilinear effect as well as its confidence intervall[25] in Figure 4 (left panel) graphically.

[24] Note that in the original argument by Bueno de Mesquita, Smith, Siverson and Morrow (2003) it was assumed that aid should decrease as a function of the size of $W$.

[25] Out of laziness we generated these confidence intervals, as well as all others reported on marginal effects below, by drawing 1000 samples from the estimated joint distribution of the coefficients. For each sample we then calculated the marginal effect for different values of $W$ and determined the 95 % intervals.

| | Original | Replication | | Replication EPR | | |
|---|---|---|---|---|---|---|
| | *Model* 1 | *Model* 2 | *Model* 3 | *Model* 4 | *Model* 5 | *Model* 6 |
| Donor coalition $W$ | $-8.275^*$ | $-8.275^*$ | $-8.397^*$ | $9.555^*$ | $9.419^*$ | $8.608^*$ |
| | (0.749) | (0.749) | (0.750) | (1.390) | (1.393) | (1.486) |
| Ln (Donor GDP p.c.) | $0.799^*$ | $0.799^*$ | $0.833^*$ | $1.066^*$ | $1.057^*$ | $1.173^*$ |
| | (0.009) | (0.009) | (0.009) | (0.011) | (0.011) | (0.013) |
| Recipient coalition $W_{t-1}$ | $-0.276$ | $-0.276$ | $-0.266$ | $-30.207^*$ | $-31.700^*$ | $-20.915^*$ |
| | (0.159) | (0.159) | (0.160) | (4.197) | (4.165) | (4.641) |
| Recipient coalition $W^2_{t-1}$ | $0.681^*$ | $0.681^*$ | $0.648^*$ | $22.712^*$ | $23.586^*$ | $17.680^*$ |
| | (0.179) | (0.179) | (0.180) | (2.430) | (2.408) | (2.694) |
| Ln(Recipient GDP p.c.)$_{t-1}$ | $3.869^*$ | $3.869^*$ | $4.148^*$ | $4.069^*$ | $4.190^*$ | $4.801^*$ |
| | (0.425) | (0.425) | (0.429) | (0.421) | (0.412) | (0.505) |
| Ln(Recipient GDP p.c.)$^2_{t-1}$ | $-0.281^*$ | $-0.281^*$ | $-0.299^*$ | $-0.290^*$ | $-0.298^*$ | $-0.330^*$ |
| | (0.027) | (0.027) | (0.027) | (0.026) | (0.026) | (0.032) |
| Recipient government share | $1.533^*$ | $1.533^*$ | $1.533^*$ | $1.167^*$ | $1.204^*$ | 1.048 |
| | (0.272) | (0.272) | (0.274) | (0.267) | (0.261) | (0.593) |
| Recipient government share$^2$ | $-0.470$ | $-0.470$ | $-0.474$ | $-0.279$ | $-0.348$ | 0.488 |
| | (0.241) | (0.241) | (0.242) | (0.237) | (0.235) | (0.868) |
| Ln(recipient population) | $1.798^*$ | $1.798^*$ | $1.740^*$ | $1.909^*$ | $1.765^*$ | $2.189^*$ |
| | (0.077) | (0.077) | (0.078) | (0.077) | (0.074) | (0.095) |
| Ln(Recipient population$^2$) | $-0.075^*$ | $-0.075^*$ | $-0.067^*$ | $-0.055^*$ | $-0.051^*$ | $-0.130^*$ |
| | (0.014) | (0.014) | (0.014) | (0.014) | (0.013) | (0.017) |
| Ln(Distance) | $-0.827^*$ | $-0.827^*$ | $-0.832^*$ | $-0.879^*$ | $-0.856^*$ | $-0.774^*$ |
| | (0.022) | (0.022) | (0.022) | (0.022) | (0.021) | (0.024) |
| Former colony | $2.618^*$ | $2.618^*$ | $2.598^*$ | $2.955^*$ | $2.962^*$ | $3.002^*$ |
| | (0.043) | (0.043) | (0.043) | (0.043) | (0.042) | (0.048) |
| US donor dummy | $0.717^*$ | $0.717^*$ | $0.618^*$ | $4.716^*$ | $4.676^*$ | $4.458^*$ |
| | (0.045) | (0.045) | (0.046) | (0.146) | (0.145) | (0.157) |
| $N$ | 39919 | 39919 | 39338 | 39338 | 40892 | 25915 |
| adj. $R^2$ | 0.478 | 0.478 | 0.480 | 0.500 | 0.496 | 0.553 |
| Resid. sd | 1.846 | 1.846 | 1.846 | 1.811 | 1.816 | 1.761 |

*Notes:* Standard errors in parentheses. $^*$ indicates significance at $p < 0.05$

The dependent variable is logged gross bilateral aid amounts. Replication based on Bueno de Mesquita and Smith (2009). Unit of analysis is the donor-recipient-year and all models comprise recipient fixed effects.

Table 6: Explaining Gross Bilateral Aid (logged), 1960-2001

In the remaining models in Table 6 we replicate Bueno de Mesquita and Smith's (2009, 328 (model 2)) analysis, first by focusing on those countries that can be covered with EPR-data (Model 3). As the results show, this has no substantive effect on the results, even though it slightly reduces the number of observations. In Model 4 we replace the original measure for $W$ with the one relying on information from the EPR data. This has as a consequence that the donor's $W$ now has a positive and statistically significant coefficient as expected by Bueno de Mesquita and Smith (2009). Regarding the effect of the recipient's $W$ we find again a curvilinear effect as in the original model, but its marginal effect is always negative for all values of $W$. This is illustrated in Figure 4 (right panel). It is worth noting that this corresponds to the effect that Bueno de Mesquita, Smith, Siverson and Morrow (2003, 483f) expected in their earlier work. This also applies to Model 5 in Table 6 for which we used all the observations available from the EPR-dataset.

As for the results obtained when only focusing on countries in which there are politically relevant ethnic groups (Model 6) these do hardly differ from those obtained for Model 4. Thus, the differences between the results obtained from models using the original measure for $W$ and those obtained with our measure appear not to be sensitive to different sample specifications.

Again, as for the previous replication, the question arises whether the results obtained based on our proposed measure are more valid than those obtained with the original measure. As neither $S$ nor the loyalty norm $W/S$ appear among the independent variables in Bueno de Mesquita and Smith's (2009, 328 (Model 2)) model, analyzing the residuals as a function of the relative values of these variables is hardly useful. Nevertheless, Table 6 provides one piece of information in favor of our proposed measure, namely the adjusted $R^2$ values for Models 3 and 4. As these values are obtained exactly on the same samples (as $R^2$ are sample-specific) the higher values obtained in Model 4 suggests that our proposed measure for $W$ provides a better fit to the data. As for the previous replication we carried out again a series of cross-validations in the same manner as above. Figure 5 shows again that the root of the mean squared prediction error is on average larger for the model using the original measures for $W$, than the one stemming from our $W$ measures. For all cross-validation based on Monte-Carlo simulations the 95% confidence interval excludes the value of 0, as it also does for the 10-fold cross-validation.[26] Consequently these cross-validations suggest that

[26]It is useful to note that for a larger number of folds the confidence intervals of the differences in

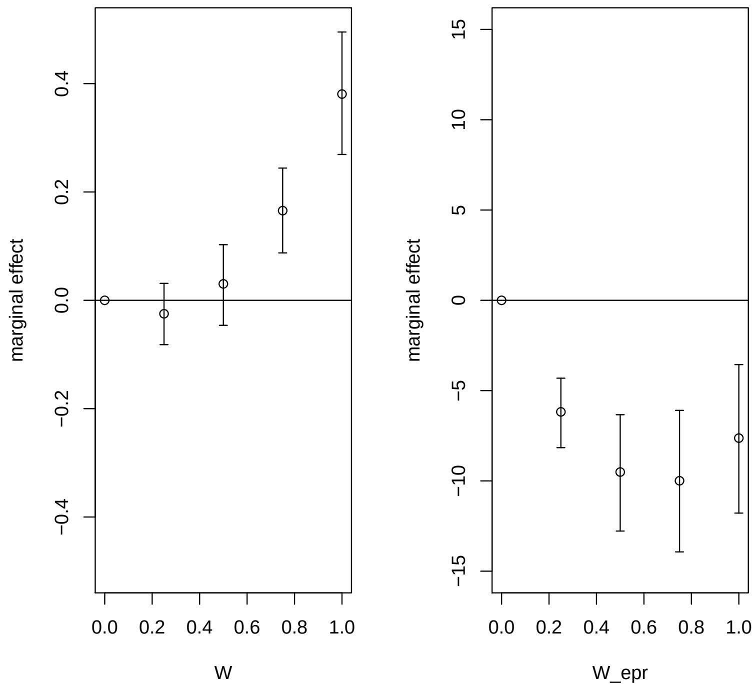


Figure 4: Curvilinear effect of recipient's $W$ on aid received

using our proposed measure for $W$ yields improved predictions.

To assess whether these findings obtained in replications of existing studies also materialize in analyses relying on a more conventional and recent empirical specification, we compiled a more up to date dataset that covers the years 1967-2009. We add the coalition size measures to a parsimonious specification that is based on widely used variables in recently published aid allocation papers (. Faye and Niehaus, 2012; Dietrich, 2013; Acht, Mahmoud and Thiele, 2015; Bermeo and Leblang, 2015; Kersting and Kilby, 2016; Bueno de Mesquita and Smith, 2016) We use information on net bilateral aid commitments from the OECD and obtained the original measures of $S$ and $W$ from the replication dataset of Smith (2016). As in the study by Bueno de Mesquita and Smith (2009) we use the $W$ of the donor country and $W$ and $W^2$ of the recipient country as additional explanatory variables. We also use the same lag structure as Bueno de Mesquita and Smith (2009) for all relevant variables and lag the

prediction errors comprise the value of 0. Consequently, the choice of the number of folds in cross-validations is not as innocuous as Ward, Greenhill and Bakke (2010, 370) seem to imply.

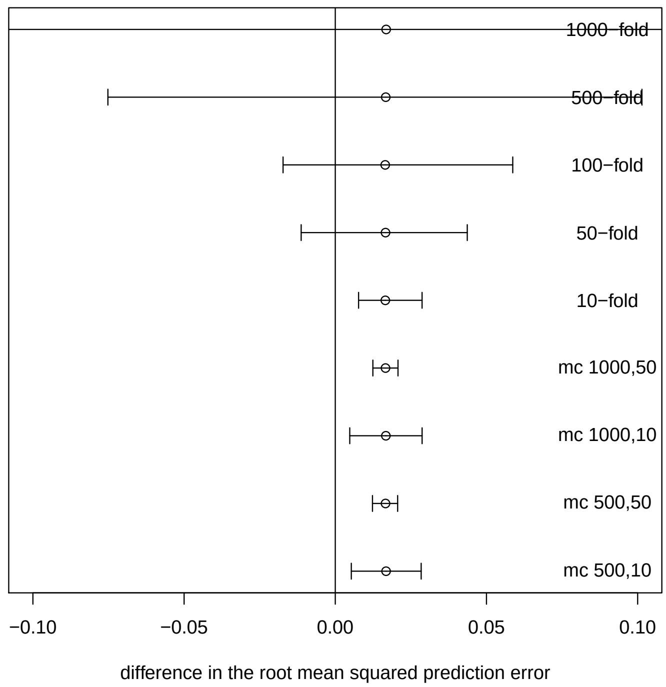


Figure 5: Comparison of prediction errors based on $k$-fold and Monte-Carlo cross-validations

remaining ones. Specifically, we include the following covariates: (logged) per capita income and (logged) population (World Bank, 2017), historic colonial relationships and geographic distance (Mayer and Zignago, 2011), and as indicator for preference alignments we use the $\kappa$ measure provided by Häge and Hug (2016), which is based on all resolution votes at the United Nations General Assembly (UNGA) (for a more detailed discussion of this measure, see Häge, 2011).

Table 7 reports the results of these estimations based on a specification with only recipient fixed effects (as in Bueno de Mesquita and Smith, 2009) (models 1 and 2) and a specification using donor, recipient and year fixed effects (models 3-6).[27] For the first two models with recipient fixed effects we find a negative (though statistically insignificant) effect of the donors' $W$ on the amount of aid, independent of using the

[27]The model also includes an indicator for whether the aid flow, in a particular year from a donor to a recipient was 0 or negative. As the dependent variable is the logged amount of aid, the log is undefined for these values and was set to 0. Adding this indicator variables assures that the functional form is not unduly affected by the coding decisions for 0 or negative values for net aid.

| | RFE, all | | D-,R-,YFE, all | | D-,R-,YFE, ethnicity relevant | |
|---|---|---|---|---|---|---|
| | $W$ | $W_{EPR}$ | $W$ | $W_{EPR}$ | $W$ | $W_{EPR}$ |
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Donor coalition $W$ | $-0.364$ | $-5.376$ | $0.772^{*}$ | $1.659^{*}$ | $1.194^{*}$ | $1.180$ |
| | $(1.446)$ | $(2.812)$ | $(0.346)$ | $(0.797)$ | $(0.451)$ | $(1.208)$ |
| Ln(Donor GDP p.c.) | $1.936^{*}$ | $2.001^{*}$ | $3.304^{*}$ | $3.312^{*}$ | $3.966^{*}$ | $3.967^{*}$ |
| | $(0.644)$ | $(0.512)$ | $(0.336)$ | $(0.336)$ | $(0.695)$ | $(0.688)$ |
| Recipient coalition $W_{t-1}$ | $0.530^{*}$ | $1.565^{*}$ | $0.298$ | $1.640^{*}$ | $0.381$ | $1.159^{*}$ |
| | $(0.189)$ | $(0.366)$ | $(0.166)$ | $(0.404)$ | $(0.241)$ | $(0.464)$ |
| Recipient coalition $W_{t-1}^2$ | $-0.337$ | $-0.842^{*}$ | $0.189$ | $-0.944^{*}$ | $0.004$ | $-0.611$ |
| | $(0.240)$ | $(0.312)$ | $(0.202)$ | $(0.344)$ | $(0.295)$ | $(0.382)$ |
| Ln(Recipient GDP p.c.)$_{t-1}$ | $2.936^{*}$ | $3.034^{*}$ | $1.818^{*}$ | $2.056^{*}$ | $1.851^{*}$ | $2.050^{*}$ |
| | $(0.585)$ | $(0.546)$ | $(0.462)$ | $(0.480)$ | $(0.621)$ | $(0.641)$ |
| Ln(Recipient GDP p.c.)$_{t-1}^2$ | $-0.245^{*}$ | $-0.248^{*}$ | $-0.154^{*}$ | $-0.168^{*}$ | $-0.156^{*}$ | $-0.168^{*}$ |
| | $(0.045)$ | $(0.041)$ | $(0.035)$ | $(0.037)$ | $(0.047)$ | $(0.048)$ |
| Ln(Recipient population) | $-1.965^{*}$ | $-1.715^{*}$ | $0.284$ | $0.187$ | $0.457$ | $0.378$ |
| | $(0.625)$ | $(0.493)$ | $(0.289)$ | $(0.295)$ | $(0.303)$ | $(0.300)$ |
| UN alignment $\kappa_{t-1}$ | $-0.520$ | $-0.204$ | $0.378^{*}$ | $0.430^{*}$ | $0.279$ | $0.306$ |
| | $(0.427)$ | $(0.350)$ | $(0.175)$ | $(0.174)$ | $(0.243)$ | $(0.240)$ |
| Former Colony | $2.703^{*}$ | $2.699^{*}$ | | | | |
| | $(0.341)$ | $(0.361)$ | | | | |
| Ln(Distance) | $-0.326$ | $-0.390$ | | | | |
| | $(0.305)$ | $(0.287)$ | | | | |
| No aid | $-1.125^{*}$ | $-1.167^{*}$ | $-0.702$ | $-0.699$ | $-0.702$ | $-0.689$ |
| | $(0.340)$ | $(0.280)$ | $(0.370)$ | $(0.368)$ | $(0.409)$ | $(0.410)$ |
| N | 43262 | 43262 | 43262 | 43262 | 27426 | 27426 |
| adj. $R^2$ | 0.142 | 0.182 | 0.017 | 0.016 | 0.009 | 0.008 |
| Resid. sd | 2.087 | 2.038 | 1.791 | 1.791 | 1.808 | 1.809 |

*Notes:* Standard errors (clustered by recipient) in parentheses. $^{*}$ indicates significance at $p < 0.05$
The dependent variable is logged net bilateral aid commitments. Unit of analysis is the donor-recipient-year and all models comprise recipient fixed effects, while the last four also comprise donor and year fixed effects.

Table 7: Explaining Net Bilateral Aid Commitments, 1967-2009

original or EPR-based measures. Similarly, we find in both cases a curvilinear effect of the recipients' $W$ which first increases and then decreases, although the decrease is statistically significant for the EPR measure only.

For the more stringent models with donor and year fixed effects added to the basic model, which removes the two time-invariant variables (former colonies and distance) (models 3 and 4) we find the theoretically expected positive effect of the donors' $W$ independent of the measure we use. For the effect of the recipients' $W$ we only find a significant curvilinear effect if we use our measure of $W$ based on the EPR data. These relationships are depicted graphically in 8. The left-hand panel nicely shows that when the original measure of $W$ is used, an increase in the values of this variable leads to a monotonic increase in aid. When our measure is used (right-hand panel) we find a

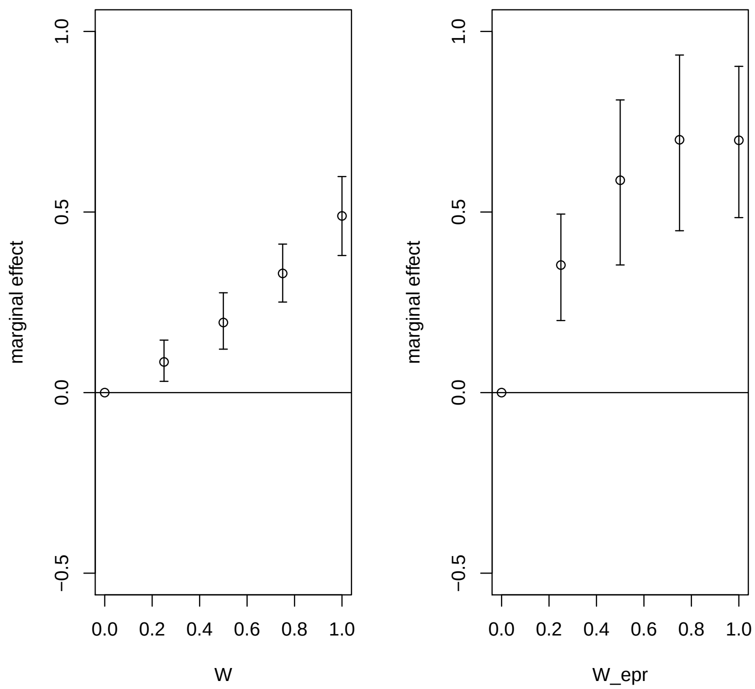


Figure 6: Curvilinear effect of recipient's $W$ on aid received

curvilinear effect, suggesting that for very high values of $W$ aid actually decreases. In line with the theory about policy concessions, UN alignment increases with the amount of aid received. Finally, the last two models (5-6) look only at those donor countries where relevant ethnic groups exist while keeping all fixed effects. These models largely confirm these findings, even though the coefficient for the donors' $W$ is no longer statistically significant.

Again, as for the replication studies we need to assess whether there is evidence suggesting that our proposed measures are preferable. The fit measures depicted in Table 7[28] suggest that our $W$s only perform better when no donor and year fixed effects are used. As soon as these additional fixed effects are introduced, the fit measures for the two sets of models are almost identical and if there is a difference, it is to the disadvantage of our measures. Nevertheless, we perform again cross-validations focusing, for simplicity's sake, on three $k$-fold cross-validations. Again we depict in Figure 7

[28]For the adjusted $R^2$ we rely on the within variance.

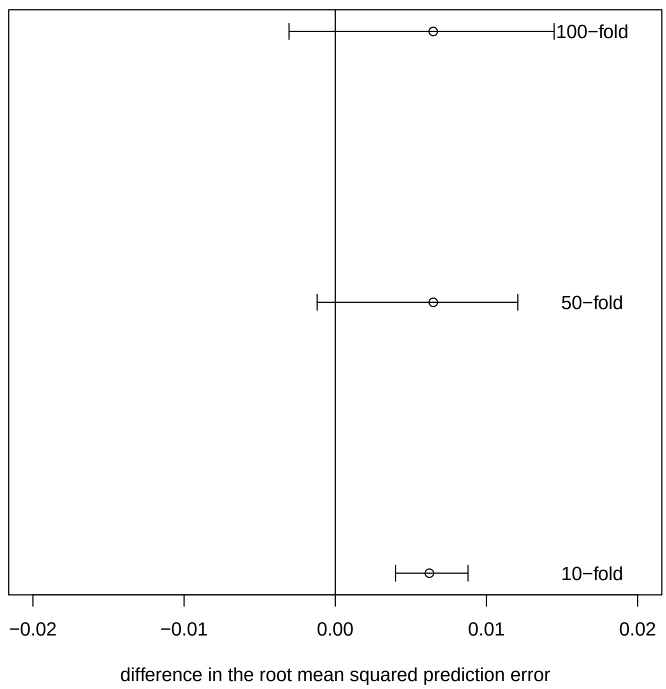


Figure 7: Comparison of prediction errors based on $k$-fold cross-validations

the difference in the root of the mean squared prediction errors between the original measures and ours. The three cross-validations show again, that on average using our measures leads to smaller root means squared prediction errors. The differences, however, are only significant for a $10$-fold cross-validation.

Finally, Table 8 reports the results of an extension based on an argument by Bueno de Mesquita and Smith (2009). More specifically, these authors argue that donors have no need to "buy policy concessions" from countries with similar preferences as theirs, while if these preferences differ too much, then the price is likely to be high.[29]

For reasons not completely transparent for the reader Bueno de Mesquita and Smith (2009) add this variable without interacting it with the other "cost" measure, namely

[29] In their study Bueno de Mesquita and Smith (2009) do find a curvilinear effect of the alignment between donor and recipient based on alliance information, but the inflection point, from an increasing effect to a decreasing effect occurs at the extreme end of the distribution of the alliance measure (which is to start with distributed in a very skewed way). More precisely in models 3 and 4 of their table 1 (Bueno de Mesquita and Smith, 2009, 328) only less than 2 %, resp. barely 3 % of the observations have such high values for the alignment measure that they lead to a decrease in aid.

| | RFE, all | | D-,R-,YFE, all | | D-,R-,YFE, ethnicity relevant | |
|---|---|---|---|---|---|---|
| | $W$ | $W_{EPR}$ | $W$ | $W_{EPR}$ | $W$ | $W_{EPR}$ |
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Donor coalition: $W$ | $-0.311$ | $-5.393$ | $0.764^*$ | $1.693^*$ | $1.182^*$ | $1.218$ |
| | $(1.450)$ | $(2.812)$ | $(0.348)$ | $(0.785)$ | $(0.450)$ | $(1.181)$ |
| Ln(Donor GDP p.c.) | $1.950^*$ | $2.021^*$ | $3.295^*$ | $3.306^*$ | $3.931^*$ | $3.941^*$ |
| | $(0.647)$ | $(0.515)$ | $(0.332)$ | $(0.333)$ | $(0.683)$ | $(0.673)$ |
| Recipient coalition $W_{t-1}$ | $0.578^*$ | $1.099^*$ | $0.343$ | $1.409^*$ | $0.453$ | $1.024$ |
| | $(0.257)$ | $(0.520)$ | $(0.233)$ | $(0.431)$ | $(0.370)$ | $(0.619)$ |
| Recipient coalition $W^2_{t-1}$ | $-0.274$ | $-0.493$ | $0.049$ | $-0.787^*$ | $-0.125$ | $-0.429$ |
| | $(0.291)$ | $(0.414)$ | $(0.247)$ | $(0.337)$ | $(0.466)$ | $(0.477)$ |
| Ln(Recipient GDP p.c.)$_{t-1}$ | $2.990^*$ | $3.165^*$ | $1.811^*$ | $2.099^*$ | $1.831^*$ | $2.135^*$ |
| | $(0.569)$ | $(0.551)$ | $(0.465)$ | $(0.497)$ | $(0.615)$ | $(0.685)$ |
| Ln(Recipient GDP p.c.)$^2_{t-1}$ | $-0.249^*$ | $-0.256^*$ | $-0.153^*$ | $-0.170^*$ | $-0.154^*$ | $-0.174^*$ |
| | $(0.044)$ | $(0.041)$ | $(0.035)$ | $(0.037)$ | $(0.045)$ | $(0.051)$ |
| Ln(Recipient population) | $-1.912^*$ | $-1.680^*$ | $0.290$ | $0.197$ | $0.474$ | $0.376$ |
| | $(0.620)$ | $(0.487)$ | $(0.279)$ | $(0.297)$ | $(0.286)$ | $(0.300)$ |
| UN alignment $\kappa_{t-1}$ | $0.787$ | $0.503$ | $0.601$ | $1.295$ | $0.725$ | $2.622$ |
| | $(0.911)$ | $(1.531)$ | $(0.606)$ | $(1.541)$ | $(0.826)$ | $(2.049)$ |
| UN alignment $\kappa^2_{t-1}$ | $-1.468$ | $-1.962$ | $-0.562$ | $-2.090$ | $-0.714$ | $-4.010$ |
| | $(0.890)$ | $(1.932)$ | $(0.684)$ | $(1.829)$ | $(0.777)$ | $(2.359)$ |
| UN alignment $\kappa_{t-1}\times$ | $1.005$ | $2.059$ | $-0.870$ | $-1.338$ | $-0.599$ | $-4.020$ |
| Recipient coalition $W_{t-1}$ | $(1.766)$ | $(4.402)$ | $(1.751)$ | $(4.885)$ | $(2.416)$ | $(6.954)$ |
| UN alignment $\kappa_{t-1}\times$ | $-2.478$ | $-1.809$ | $1.398$ | $0.781$ | $0.634$ | $1.781$ |
| Recipient coalition $W^2_{t-1}$ | $(2.542)$ | $(3.509)$ | $(1.932)$ | $(3.838)$ | $(2.712)$ | $(5.704)$ |
| No aid | $-1.069^*$ | $-1.180^*$ | $-1.135$ | $-0.706$ | $-0.705$ | $-0.699$ |
| | $(0.340)$ | $(0.279)$ | $(0.368)$ | $(0.368)$ | $(0.408)$ | $(0.416)$ |
| UN alignment $\kappa^2_{t-1}\times$ | $-2.439$ | $0.632$ | $1.165$ | $4.400$ | $0.200$ | $9.228$ |
| Recipient coalition $W_{t-1}$ | $(2.726)$ | $(5.972)$ | $(2.188)$ | $(6.483)$ | $(2.711)$ | $(9.334)$ |
| UN alignment $\kappa^2_{t-1}\times$ | $4.101$ | $-0.024$ | $-1.199$ | $-2.692$ | $0.357$ | $-5.468$ |
| Recipient coalition $W^2_{t-1}$ | $(3.820)$ | $(4.869)$ | $(2.388)$ | $(5.206)$ | $(3.164)$ | $(7.928)$ |
| N | 43262 | 43262 | 43262 | 43262 | 27426 | 27426 |
| adj. $R^2$ | $0.145$ | $0.183$ | $0.017$ | $0.017$ | $0.009$ | $0.009$ |
| Resid. sd | $2.086$ | $2.037$ | $1.791$ | $1.791$ | $1.808$ | $1.808$ |

*Notes:* Standard errors (clustered by recipient) in parentheses. $^*$ indicates significance at $p < 0.05$
The dependent variable is logged bilateral aid amounts. Unit of analysis is the donor-recipient-year and all models comprise recipient fixed effects, while the last four also comprise donor and year fixed effects.

Table 8: Explaining Gross Aid

the recipient's $W$. As both measures are assumed to have curvilinear effects, omitting interaction effects is likely to lead to a misspecification of the model. For this reason we use not only our alignment measure and its square, but also the interactions of these variables with $W$ and $W^2$ as additional regressors to assess how preference alignments affect the effect of $W$

The results depicted in Table 8 show very similar results for the donors' $W$ as those reported in Table 7, namely that the positive effect materializes only with donor fixed effects. Regarding the recipients' $W$ the coefficients are more difficult to interpret because of the interaction with the UN alignment measure. For this reason we depict in Figure 8 how for the different values of $\kappa$ (i.e., the extent to which there is preference alignment between donor and recipient) the marginal effect of the recipients' $W$ changes. Figure 8 nicely shows that increasing voting alignment depresses the marginal effect of $W$. The latter (left-hand panel) increases almost linearly if the voting alignment is at its lowest value. As voting alignment increases, the marginal effect becomes negative, especially so for larger values of $W$. For the highest value of $W$ the confidence intervals for the curves linked to the two extreme values for $\kappa$, namely -0.01 and 1 fail to overlap.[30] For our measure of $W$ (right-hand panel) we find again the curvilinear effect of $W$. For low values of our voting alignment measure we find a monotonic and positive relationship, implying that "bitter enemies" only get aid if their winning coalition is large. As preference alignment increases, the marginal effect of $W$ becomes negative, however, it remains negative (and increasing in absolute value) over the full range of $W$. Consequently, these results seem to suggest that "friends" are less rewarded as their $W$ increases.

Finally, for completeness' sake we depict in Figure 9 three cross-validations based on models 3 and 4 from Table 8. The figure underlines what we have seen in Figure 7, namely that our measure leads to smaller root mean squared prediction errors, and these differences are statistically significant if we consider 10-fold cross-validations.

[30]Needless to say, that for some smaller values of $W$ there is still a significant difference, as the overlap of the 95% confidence intervals does not imply statistically insignificant differences.

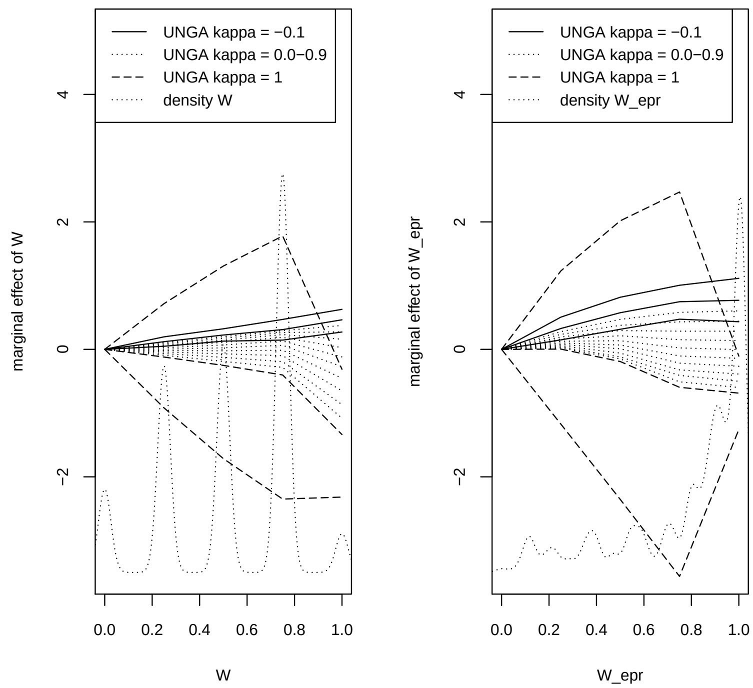


Figure 8: Marginal effect of recipient's $W$ as a function of UN alignment ($\kappa$) on aid received

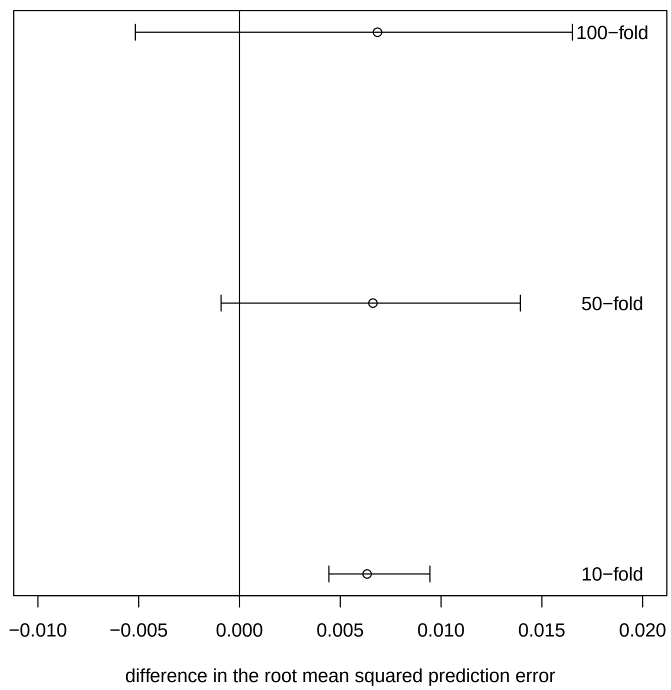


Figure 9: Comparison of prediction errors based on $k$-fold cross-validations

# 5 Discussion and Conclusion

The measures used by Bueno de Mesquita, Smith, Siverson and Morrow (2003) for their central concepts of selectorate and winning coalition have been criticized by multiple scholars. In this paper, we proposed alternative measurements of $W$ and $S$ which fulfill three conditions: (i) they span both democracies and dictatorships by offering continuous measures, and (ii) they avoid unrealistic cases in which $W > S$ or (iii) where $S = 0$. We believe that our measure succeeds on all three counts. Although no objectively valid way exists that would decide which measure is more appropriate, our empirical tests based on models of foreign aid allocation derived from the selectorate theory demonstrate the usefulness of our indicators. In both replications, when focusing on the observations for which both proposed measures can be derived, our measures lead to a better fit of the model to the data. In addition, cross-validations provide consistent evidence in favor of our proposed measures, even though the differences fail to reach statistical significance for one of the analyses, namely the one with fewer observations and a more complicated empirical model.

Nevertheless some caveats remain. First, our measures presume that ethnic divisions and processes of exclusion and discrimination underlie political processes. Our measure will perform less well where non-ethnic cleavages predominate. Second, our actor-centered measure of $W$ and $S$ does not comply with Bueno de Mesquita, Smith, Siverson and Morrow's (2003) demand for institution-based measures. Nevertheless, Bormann, Cederman, Gates, Graham, Hug, Strøm and Wucherpfennig (2014) demonstrate that formal power-sharing institutions compiled by Strøm, Gates, Graham and Strand (2017) considerably affect EPR-based measures of ethnic coalitions.[31] Finally, while we obtain similar findings for some analyses as Bueno de Mesquita, Smith, Siverson and Morrow (2003) and Bueno de Mesquita and Smith (2009), we do not reproduce all their findings. Some of our results are more in line with the theoretical predictions, while others contradict them. Nevertheless, we believe that the empirical measures of the size of the selectorate and the winning coalitions introduced in this paper improve on earlier efforts.

[31] Bormann, Cederman, Gates, Graham, Hug, Strøm and Wucherpfennig (2014) note also, however, that formal institutions are not the only conduit through which ethnic power-sharing can come about. Thus, their analysis reflects in part the fact that non-institutional aspects of authoritarian politics as highlighted by Gallagher and Hanson (2015) still play a role (see also Roessler, 2011; Roessler, 2016; Roessler and Ohls, 2017 (forthcoming)).

# Appendix

In table 9 we list the countries covered in the analysis underlying table 4 in the main text, while table 10 reports the descriptive statistics of the data underlying the analyses reported in tables 7 and 8.

| country | n original | n replication |
|---|---|---|
| Albania | 4 | 4 |
| Algeria | 5 | 5 |
| Argentina | 4 | 4 |
| Australia | 26 | 26 |
| Austria | 5 | 5 |
| Bahamas | 26 | 0 |
| Bahrain | 18 | 9 |
| Bangladesh | 1 | 1 |
| Barbados | 20 | 0 |
| Belarus | 7 | 7 |
| Belgium | 29 | 29 |
| Bhutan | 11 | 11 |
| Bolivia | 6 | 6 |
| Botswana | 20 | 20 |
| Brazil | 7 | 7 |
| Bulgaria | 1 | 1 |
| Burkna Faso | 4 | 4 |
| Burundi | 9 | 9 |
| Cameroon | 13 | 13 |
| Canada | 27 | 27 |
| Ceylon | 1 | 1 |
| Chad | 4 | 4 |
| Chile | 13 | 13 |
| Colombia | 6 | 6 |
| Comoro IS | 1 | 0 |
| Congo | 5 | 5 |
| Congo DR | 1 | 1 |
| Congo Rep | 2 | 2 |
| Costa Rica | 13 | 13 |
| Cote D'ivor | 6 | 6 |
| Cyprus | 8 | 8 |
| Czech Rep | 7 | 7 |
| Denmark | 8 | 8 |
| Domin Rep | 2 | 2 |
| Egypt | 2 | 2 |
| El Salvador | 3 | 3 |
| Estonia | 4 | 4 |
| Eth'pia PDR | 6 | 6 |
| Ethiopia | 6 | 6 |
| Fiji | 21 | 21 |
| Finland | 28 | 28 |
| Gambia | 10 | 10 |
| Georgia | 3 | 3 |
| Germany | 9 | 0 |
| Ghana | 17 | 17 |
| Greece | 19 | 19 |
| Guatemala | 1 | 1 |
| Guyana | 16 | 16 |
| Haiti | 4 | 4 |
| Hungary | 16 | 16 |
| Iceland | 27 | 0 |
| India | 26 | 26 |
| Indonesia | 28 | 28 |
| Ireland | 3 | 3 |
| Israel | 28 | 28 |
| Italy | 13 | 13 |
| Jamaica | 6 | 6 |
| Japan | 4 | 4 |
| Jordan | 25 | 25 |
| Kazakhstan | 3 | 3 |
| Kenya | 7 | 7 |
| Korea Rep | 28 | 28 |
| Kyrgyzstan | 1 | 1 |
| Latvia | 6 | 6 |
| Lebanon | 6 | 6 |
| Lesotho | 9 | 9 |
| Lithuania | 3 | 3 |
| Luxembourg | 25 | 0 |
| Madagascar | 8 | 8 |
| Malagasyr | 4 | 4 |
| Malawi | 14 | 14 |
| Malaysia | 15 | 15 |
| Maldive IS | 14 | 0 |
| Mali | 4 | 4 |
| Malta | 26 | 0 |

| | | |
|---|---|---|
| Mauritius | 28 | 28 |
| Mexico | 26 | 26 |
| Moldova | 5 | 5 |
| Mongolia | 7 | 7 |
| Morocco | 25 | 25 |
| Namibia | 4 | 0 |
| Nauru | 1 | 0 |
| Nepal | 19 | 19 |
| Netherlands | 12 | 12 |
| New Zealand | 29 | 29 |
| Nicaragua | 10 | 10 |
| Nigeria | 5 | 5 |
| Norway | 28 | 28 |
| Oman | 25 | 25 |
| Pakistan | 25 | 25 |
| Panama | 7 | 7 |
| Papua New G | 25 | 25 |
| Paraguay | 11 | 11 |
| Peru | 6 | 6 |
| Philippines | 28 | 28 |
| Poland | 6 | 6 |
| Portugal | 6 | 6 |
| Rhodesia | 4 | 4 |
| Romania | 1 | 1 |
| Russian Fed | 2 | 2 |
| Rwanda | 9 | 9 |
| Senegal | 10 | 10 |
| Seychelles | 4 | 0 |
| Sierra Leo | 23 | 23 |
| Singapore | 29 | 29 |
| Slovak Rep | 4 | 4 |
| Slovenia | 7 | 7 |
| So Africa | 12 | 12 |
| Solomon Is | 7 | 0 |
| Spain | 26 | 26 |
| Sri Lanka | 28 | 28 |
| St Vincent | 14 | 0 |
| Suriname | 6 | 0 |
| Swaziland | 12 | 11 |
| Sweden | 12 | 12 |
| Switzerland | 23 | 23 |
| Tajikistan | 1 | 1 |
| Thailand | 29 | 29 |
| Togo | 5 | 5 |
| Trinidad | 3 | 3 |
| Tunisia | 29 | 29 |
| Turkey | 27 | 27 |
| Uganda | 1 | 1 |
| UK | 28 | 28 |
| Uruguay | 23 | 23 |
| US | 28 | 28 |
| Vanuatu | 12 | 0 |
| Venezuela | 17 | 17 |
| Westn Samoa | 1 | 0 |
| Zaire | 25 | 25 |
| Zambia | 10 | 10 |
| Zimbabwe | 18 | 18 |
| total | 1646 | 1439 |

Table 9: Countries covered in analysis on Who get's aid . . . ?

| | Mean | Variance | Skewness | Kurtosis | Min | Max | N |
|---|---|---|---|---|---|---|---|
| Ln (aid) | 1.37 | 5.92 | -0.18 | -0.56 | -4.61 | 9.91 | 43262 |
| Donor coalition $W$ | 0.99 | 0.00 | -6.80 | 59.97 | 0.25 | 1.00 | 43262 |
| Ln(Donor GDP p.c.) | 10.46 | 0.12 | -0.13 | 1.29 | 9.17 | 11.63 | 43262 |
| Recipient coalition$_{t-1}$ | 0.53 | 0.07 | -0.43 | -0.81 | 0.00 | 1.00 | 43262 |
| Recipient coalition$^2_{t-1}$ | 0.35 | 0.07 | 0.40 | -0.53 | 0.00 | 1.00 | 43262 |
| Ln(Recipient GDP p.c.)$_{t-1}$ | 7.35 | 1.20 | 0.23 | -0.62 | 4.81 | 11.65 | 43262 |
| Ln(Recipient GDP p.c.)$^2_{t-1}$ | 55.24 | 271.00 | 0.53 | -0.20 | 23.12 | 135.79 | 43262 |
| Ln(Population) | 16.28 | 2.33 | 0.54 | 0.53 | 12.83 | 21.01 | 43262 |
| UN alignment $(\kappa)_{t-1}$ | 0.18 | 0.06 | 1.55 | 1.36 | -0.09 | 1.00 | 43262 |
| Donor coalition$_{epr_{t-1}}$ | 0.96 | 0.01 | -2.88 | 7.30 | 0.57 | 1.00 | 43262 |
| Recipient coalition$_{epr_{t-1}}$ | 0.79 | 0.06 | -1.26 | 0.50 | 0.02 | 1.00 | 43262 |
| Recipient coalition$^2_{epr_{t-1}}$ | 0.69 | 0.10 | -0.79 | -0.72 | 0.00 | 1.00 | 43262 |

Table 10: Descriptive statistics

# References


Acht, Martin, Toman Omar Mahmoud and Rainer Thiele. 2015. "Corrupt governments do not receive more state-to-state aid: Governance and the delivery of foreign aid through non-state actors." *Journal of Development Economics* 114:20–33.

Alesina, Alberto, Stelios Michalopoulos and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy* 124(2):428–488.

Alvarez, Mike, Jose Antonio Cheibub, Fernando Limongi and Adam Przeworski. 1996. "Classifying Political Regimes." *Studies in Comparative International Development* 31(2 Summer):3–36.

Arlot, Sylvain and Alain Celisse. 2010. "A survey of cross-validation procedures for model selection." *Statistics surveys* 4:40–79.

Baldwin, Kate and John D. Huber. 2010. "Economic versus cultural differences: Forms of ethnic diversity and public goods provision." *American Political Science Review* 104(4):644–662.

Banks, Arthur S., Alan J. Day and Thomas C. Mueller, eds. 1996. *Political Handbook of the World*. CSA Publications.

Baskaran, Thushyanthan, Yogesh Uppal and Brian Min. 2015. "Election cycles and electricity provision: evidence from a quasi-experiment with Indian special elections,." *Journal of Public Economics* 126:64–73.

Bates, Robert H. 1974. "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* 6(4):457–484.

Beiser, Janina and Nils W Metternich. 2016. "Ethnic exclusion and the logic of political survival." Paper prepared for presentation at the ENCoRe Workshop, Geneva (January 21-22).

Bermeo, Sarah B. and David Leblang. 2015. "Migration and Foreign Aid." *International Organization* 69(3):627657.

Birnir, Jhanna, Jonathan Wilkenfeld, James Fearon, David Laitin, Ted Robert Gurr, Dawn Brancati, Stephen Saideman, Amy Pate and Agatha Hultquist. 2015. "Socially Relevant Ethnic Groups, Ethnic Structure and AMAR." *Journal of Peace Research* 52(1):110–115.

Bormann, Nils-Christian. 2014. "The Causes and Consequences of Ethnic Power-Sharing." Dissertation ETH Zurich.

Bormann, Nils-Christian, Lars-Erik Cederman, Scott Gates, Benjamin A. T. Graham, Simon Hug, Kaare Strøm and Julian Wucherpfennig. 2014. "Does formal or informal power sharing produce peace?" Paper prepared for presentation at the Annual Conference of the European Political Science Association Edinburgh, June, 2014.

Briggs, Ryan C. 2012. "Electrifying the base? Aid and incumbent advantage in Ghana." *The Journal of Modern African Studies* 50(04):603–624.

Briggs, Ryan C. 2014. "Aiding and Abetting: Project Aid and Ethnic Politics in Kenya." *World Development* 64:194–205.

Brown, Graham K., Frances Stewart and Arnim Langer. 2010. The Implications of Horizontal Inequality for Aid. In *Foreign Aid for Development*, ed. Mavrotas George. Oxford: Oxford University Press pp. 94–121.

Bruk, Salomon I. 1964. *Atlas Narodov Mira*. Moscow: Academy of Science USSR.

Bueno de Mesquita, Bruce and Alastair Smith. 2007. "Foreign aid and policy concessions." *Journal of Conflict Resolution* 51(2):251–284.

Bueno de Mesquita, Bruce and Alastair Smith. 2009. "A Political Economy of aid." *International Organization* 63(2):309–340.

Bueno de Mesquita, Bruce and Alastair Smith. 2016. "Competition and Collaboration in Aid-for-Policy Deals." *International Studies Quarterly* 60(3):413–426.

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

Bueno de Mesquita, Bruce, George Downs and Alastair Smith. 2017 (forthcoming). "A Political Economy of Income Tax Policies." *Political Science Research and Methods* .

Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93:791–807.

Bueno de Mesquita, Bruce and Randolph Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review* 89:841–855.

Burgess, Robin, Rémi Jedwab, Edward Miguel, Ameet Morjaria and Gerard Padro i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review* 105(6):1817–1851.

Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic-Nationalist Groups Rebel: New Data and Analysis." *World Politics* 62(1):87–119.

Cederman, Lars-Erik, Kristian Skrede Gleditsch and Halvard Buhaug. 2013. *Inequality, Grievances and Civil War*. Cambridge University Press.

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

Clarke, Kevin A. and Randall W. Stone. 2008. "Democracy and the Logic of Political Survival." *American Political Science Review* 102(3):387–392.

Dietrich, Simone. 2013. "Bypass or Engage? Explaining Donor Delivery Tactics in Foreign Aid Allocation." *International Studies Quarterly* 57(4):698–712.

Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley Parks, Paul A. Raschky and Michael J. Tierney. 2016. "Aid on Demand: African Leaders and the Geography of China's Foreign Assistance." AidData Working Paper No. 3 (revised).

Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–1250.

Efron, Bradley. 1983. "Estimating the error rate of a prediction rule: improvement on cross-validation." *Journal of the American Statistical Association* 78(382):316–331.

Faye, Michael and Paul Niehaus. 2012. "Political Aid Cycles." *American Economic Review* 102(7):3516–30.

Francois, Patrick, Ilia Rainer and Francesco Trebbi. 2015. "How Is Power Shared In Africa?" *Econometrica* 83(2):465–503.

Fuchs, Andreas, Axel Dreher and Peter Nunnenkamp. 2014. "Determinants of Donor Generosity: A survey of the aid budget literature." *World Development* 56:172–199.

Gallagher, Mary and Jonathan Hanson. 2013. Authoritarian survival, resilience, and the selectorate theory. In *Why Communism Did Not Collapse: Understanding Authoritarian Regime Resilience in Asia and Europe*, ed. Martin Dimitrov. Cambridge Univerisity Press pp. 185–204.

Gallagher, Mary and Jonathan Hanson. 2015. "Power Tool or Dull Blade? Selectorate Theory for Autocracies." *Annual Review of Political Science* 18(1):367–385.

Geisser, Seymour. 1975. "The predictive sample reuse method with applications." *Journal of the American Statistical Association* 70(350):320–328.

Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, NY: Cornell University Press. 12.

Gurr, Ted R. 2000. *Peoples versus States: Minorities At Risk in the New Century*. Washington, DC: United States Institute of Peace Press.

Häge, Frank M. 2011. "Choice or circumstance? Adjusting measures of foreign policy similarity for chance agreement." *Political Analysis* 19(3):287–305.

Häge, Frank M. and Simon Hug. 2016. "Consensus voting and similarity measures in IOs." *International Interactions* 42(3):503–529.

Hechter, Michael. 2000. *Containing Nationalism*. New York, NY: Oxford University Press.

Heger, Lindsay and Idean Salehyan. 2007. "Ruthless rulers: Coalition size and the severity of civil conflict." *International Studies Quarterly* 51(2):385–403.

Jablonski, Ryan S. 2014. "How aid targets votes: the impact of electoral incentives on foreign aid distribution." *World Politics* 66(02):293–330.

Kasara, Kimuli. 2007. “Tax Me if You Can: Ethnic Geography, Democracy and the Taxation of Agriculture in Africa.” *American Political Science Review* 101(1):159–172.

Kennedy, R. 2009. “Survival and accountability: an analysis of the empirical support for selectorate theory.” *International Studies Quarterly* 53(3):695–714.

Kersting, Erasmus and Christopher Kilby. 2016. “With a little help from my friends: Global electioneering and World Bank lending.” *Journal of Development Economics* 121:153 – 165.

Kroth, Verena, Valentino Larcinese and Joachim Wehner. 2016. “A Better Life for All? Democratization and Electrification in Post-Apartheid South Africa.” *The Journal of Politics* 78(3):774–791.

Mann, Michael. 2005. *The Dark Side of Democracy*. Cambridge, UK: Cambridge University Press. 8.

Marcum, Anthony S. and Jonathan N. Brown. 2016. “Overthrowing the Loyalty Norm: The Prevalence and Success of Coups in Small-coalition Systems, 1950 to 1999.” *Journal of Conflict Resolution* 60(2):256–282.

Marshall, Monty G., Ted Gurr, Christian Davenport and Keith Jaggers. 2002. “Polity IV, 1800-1999: A Reply to Munck and Verkuilen.” *Comparative Political Studies* 35(1):40–45.

Mayer, Thierry and Soledad Zignago. 2011. “Notes on CEPII’s Distances Measures: the GeoDist Database.” CEPII Working Paper 2011-25.

Min, Brian. 2015. *Power and the Vote: Elections and Electricity in the Developing World*. Cambridge University Press.

Morelli, Massimo and Dominic Rohner. 2015. “Resource Concentration and Civil Wars.” *Journal of Development Economics* 117:32–47.

Morrow, James D, Bruce Bueno de Mesquita, Randolph M. Siverson and Alastair Smith. 2008. “Retesting selectorate theory: Separating the effects of W from other elements of democracy.” *American Political Science Review* 102(3):393–400.

Pepinsky, Thomas. 2014. “The Institutional Turn in Comparative Authoritarianism.” *British Journal of Political Science* 44(3):631–53.

Powell, G. Bingham. 2000. *Elections as Instruments of Democracy. Majoritarian and Proportional Visions*. New Haven: Yale University Press.

Roessler, Philip. 2011. “The Enemy Within: Personal Rule, Coups and Civil War in Africa.” *World Politics* 63(2):300–346.

Roessler, Philip. 2016. *Ethnic Politics and State Power in Africa: The Logic of the Coup-Civil War Trap*. Cambridge University Press.

Roessler, Philip and David Ohls. 2017 (forthcoming). “The Societal Origins of Powersharing. Ethnic Geography, Threat Capabilities and Horizontal Accountability in Weak States.” *International Organization* .

Slater, Dan. 2010. *Ordering Power: Contentious Politics and Authoritarian Leviathans in Southeast Asia*. Cambridge, MA: Cambridge University Press. 55.

Smith, Alastair. 2016. "Leader turnover, institutions, and voting at the UN General Assembly." *Journal of Conflict Resolution* 60(1):143–163.

Smith, Alastair and Raymond R. Vreeland. 2010. The Survival of Political Leaders and the IMF Programs. In *The Impact of the IMF and the World Bank*, ed. Gustav Ranis, James R. Vreeland and Stephen Kosack. New York: Routledge pp. 263–289.

Stewart, Frances. 2008. Horizontal Inequalities and Conflict: An Introduction and Some Hypotheses. In *Horizontal Inequalities and Conflict*, ed. Frances Stewart. New York: Palgrave Macmillan.

Strøm, Kaare, Scott Gates, Benjamin A.T. Graham and Håvard Strand. 2017. "Inclusion, Dispersion, and Constraint: Powersharing in the World's States, 1975-2010." *British Journal of Political Science* 47(1):165–185.

van der Veen, A. Maurits and David D. Laitin. 2012. Modeling the Evolution of Ethnic Identity. In *Constructivist Theories of Ethnic Polictics*, ed. Kanchan Chandra. New York, NY: Oxford University Press pp. 277–311.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family." *Journal of Conflict Resolution* 59(7):1327–1342.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by P-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47(5):363–375.

Weidmann, Nils B., Suso Benitez-Baleato, Philipp Hunziker, Eduard Glatz and Xenofontas Dimitropoulos. 2016. "Digital discrimination: Political bias in Internet service provision across ethnic groups." *Science* 353(6304):1151–1155.

Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflict: Shadows of Modernity*. Cambridge, UK: Cambridge University Press.

World Bank. 2017. *World Development Indicators*. World Bank Publications, Washington, D.C., United States.

Wright, Joseph and Matthew Winters. 2010. "The Politics of Effective Foreign Aid." *Annual Review of Political Science* 13:61–80.