Article:

Heinrich, T and Kobayashi, Y orcid.org/0000-0003-3908-1074 (2022) Evaluating explanations for poverty selectivity in foreign aid. Kyklos, 75 (1). pp. 30-47. ISSN 0023-5962

https://doi.org/10.1111/kykl.12284

Evaluating Explanations for Poverty Selectivity in Foreign Aid

Tobias Heinrich* Yoshiharu Kobayashi†

June 16, 2021‡

*Associate Professor of Political Science, Department of Political Science, University of South Carolina, 349 Gambrell Hall, 817 Henderson Street, Columbia, South Carolina, USA, 29208, E-mail: heinrict@mailbox.sc.edu

†Associate Professor of Global Political Economy and Development, School of Politics and International Studies, University of Leeds, Social Sciences Building, Leeds, UK, LS2 9JT, E-mail: y.kobayashi@leeds.ac.uk

‡We are thankful for useful feedback and suggestions from the referees, the *Kyklos* editor, Faisal Ahmed, David Bearce, Mwita Chacha, Marina Magaldi, Carla Martinez Machain, Elena McLean, Cliff Morgan, Tim Peterson, Liana Reyes-Reardon, Alastair Smith, Harvey Starr, and Spencer Willardson as well as from participants at presentations at Rice University, the Centro de Investigación y Docencia Económicas (CIDE), Nazarbayev University, University College London, the University of Houston, the University of Mannheim, the 2016 International Studies Association Annual Convention, and 2016 Midwest Annual Convention. Leah Long provided research assistance. Replication files are provided in the supplementary material and at `www.theinrich.net/research`.

Abstract

Ending global poverty has been at the forefront of the development agenda since the 1970s, but many donors have failed to target their funds toward this goal. Activists have tackled this issue by appealing to donors' humanitarian motives, but we know little about what explains donors' decisions on how much to give to the poorest countries. This paper develops the *donor motivation* and *foreign policy* approaches and identify donors' development motives and their budget sizes as potential determinants of poverty selectivity. We evaluate their explanatory power by assessing whether their relationships with selectivity are in the hypothesized directions and generalize beyond a particular dataset. Employing cross-validation and Bayesian Model Averaging, we find few measures of donor motivations provide a generalizable and hypothesized explanation for poverty selectivity. In contrast, donor budget sizes exhibit a relationship that is both hypothesized and externally valid. Our study offers the first systematic analysis of aid selectivity and generates implications for recent approaches to improve the quality of foreign aid and the conventional approach to study foreign aid allocation and donor motives.

1 Introduction

Ending global poverty has long been front and center on the agenda of international development community (Finnemore, 1997; Riddell, 2007). It is the core of major global development initiatives such as the Sustainable Development Goals and is explicitly identified in mission statements of many donors of foreign aid.[1,2] A policy step toward achieving this objective has been to target scarce aid funds toward the very poorest countries. While some donors appear committed to directing their aid to the poorest countries, many have failed to even take initial steps (Dollar and Levin, 2006; Easterly, 2007).[3]

Activists have been castigating donors for their failures to act. One prominent approach is to measure donors' 'poverty selectivity'—the share of aid going to the poorest countries—and to publicly name and shame poor allocation choices (Roodman, 2012; Knack et al., 2011; Easterly and Williamson, 2011). Naming the inadequate performers can generate focal points for citizens, non-governmental organizations, and international organizations to criticize deficient aid allocations and shame donors into better targeting their aid.[4,5]

While the operative consensus centers on increasing poverty selectivity,[6] little is known about what explains donors' concentration of aid on the poorest countries. There is a rich

1 For example, the official goal of the World Bank is to achieve 'a world free of poverty.' The United States Agency for International Development's mission statement declares, 'We partner to end extreme poverty and promote resilient, democratic societies while advancing our security and prosperity.'

2 We use 'foreign aid' and 'aid' to refer to development aid, which excludes military aid.

3 Also see Alesina and Dollar (2000, p. 47-49), Nunnenkamp and Thiele (2006, p. 1182-1184), and Collier and Dollar (2002, p. 1495-1497). Briggs (2017, 2018) shows that not even the poorest within each country are targeted.

4 On such informational effects, see Murdie and Davis (2012), Kelley and Simmons (2015), and Honig and Weaver (2019).

5 There are two approaches to these rank-name-shame efforts. The first common approach is the creation of selectivity indices McGillivray (1989, p. 562-563), which are the weighted sum of a donor's aid flows to all recipients where the weights represent 'recipient need' (i.e. recipients' GDP per capita) (Roodman, 2012; Clist, 2011; Easterly and Williamson, 2011). The second is to use regressions to model donors' aid allocations as a function of indicators of poverty (Dudley and Montmarquette, 1976; Maizels and Nissanke, 1984). The estimated coefficients on these 'recipient need' variables are said to indicate how responsive donors are to recipients 'in need' and therefore are interpreted as the level of donors' selectivity (Knack et al., 2011; Easterly, 2007; Dollar and Levin, 2006).

6 Poverty selectivity is not the only aspect of this 'quality of aid' agenda (Knack et al., 2011; Easterly and Williamson, 2011). Other foci include reducing donors' overhead costs, greater transparency on aid projects and delivery channels, selectivity by regime type and corruption levels, etc. While we will return to these at the end, this paper centers on poverty selectivity as it has been a commonly accepted principle since the early days of development assistance and is the only criterion explicitly targeting the goal of ending extreme poverty. Other metrics aim to make aid more effective anywhere and thus might actually detract from ending extreme poverty.

literature examining how donors allocate aid among recipient countries, which has produced a few stylized facts, including that Nordic countries target more of their aid to the poorest countries than large donors such as the United States and Japan, and that U.S. aid allocations became more selective at the end of the Cold War.[7] These studies attribute the observed behavior to differences in unobservable donors' humanitarian motivations. While this approach is sensible, its validity is based on an assumption that a donor's motives behind aid are related to the poverty selectivity of its aid allocation, which has been rarely examined.

In this paper, we begin by developing two theoretical approaches, each of which focuses on specific aspects of donors, to identify potential determinants of poverty selectivity. First, we develop an approach that centers on donors' *motivations*. In this framework, donors' aid allocation priorities are a direct translation of donors' intentions behind foreign aid.[8] The perspective points to donor motives as a key determinant of poverty selectivity: donors that care more about international development and the poorest elsewhere in the world would direct a greater deal of its aid to the poorest countries.

The second approach draws on a broader foreign policy literature and sees foreign aid as one of donors' policy tools to bring about changes in the international system (Clark et al., 2008; Palmer and Morgan, 2011; Bueno de Mesquita and Smith, 2009). This *foreign policy* perspective differs from the motivation approach in that it assumes *all* donors are motivated by their 'selfish' concerns. In order to receive aid, a country must agree to change its policies to the donor's liking.[9] These targeted policies can be regarding the recipient's development (like bureaucracy reform) or some policy concession that donors like for selfish reasons (e.g. votes in the United Nations). A key implication is that in the eyes of the donor, poorer countries are more amenable to change their policies for smaller amounts of aid and are thus less expensive. Consequently, a donor with a smaller governmental budget[10] can afford to seek changes only from the poorest recipients, but as its budget increases, it can effect policy change in richer recipients. This logic points to

7 See McGillivray (1989, p. 563-565), Knack, Rogers, & Eubank (2011, p. 1912-1913), and Easterly & Williamson (2011, p. 1938-1939), Easterly (2007).

8 The argument are largely an imputation of the theoretical underpinnings behind the ranking efforts (Roodman, 2012; Knack et al., 2011; Easterly and Williamson, 2011) and empirical research to identify motives behind aid giving (McKinley and Little, 1979; Alesina and Dollar, 2000).

9 This assumption is commonplace in the literature; see Wang (2016), Bueno de Mesquita and Smith (2009), Heinrich (2013), Bermeo (2017), and Carter and Stone (2015).

10 We use 'budget' and 'resources' interchangeably when they describe the total governmental expenditure that the donor possesses to produce domestic and foreign policies. Note that this is not the size of the aid budget, not scaled as a percentage of the donor's economy, and also does not refer to 'latent power' (e.g. iron and steel production).

donors' budgetary resources as a potential determinant of poverty selectivity: as donors' government budget levels become smaller, donors concentrate a greater *share* of their aid on the poorest recipients.

We empirically assess which of the two features of donors derived from the two distinct theoretical approaches—their motivations and budget sizes—contribute more explanatory power to their poverty selectivity. Using a data set on 25 OECD Development Assistance Committee (DAC) donors between 2001 and 2019, we consider whether coefficient estimates on indicators of donor motives and resources are in the hypothesized directions as well as whether the relationships generalize beyond the particular sets of data. The former is an in-sample analysis typically pursued by scholars of foreign aid while the latter is assessed by analyzing the out-of-sample predictive accuracy. The use of out-of-sample fit helps us avoid a danger of overfitting and drawing conclusions based on some idiosyncratic features of a particular dataset.[11]

While measuring the size of governmental budgets is straightforward, finding proxies for donors' motivations behind is surely not. We searched the vast existing aid allocation literature for arguments that variables capture a donor's motivations behind the provision of aid. These proxies include gender and ideological influences in donor-side politics generally and in aid specifically, whether the aid agency is independent, the extent of social global connections, among several others. As we have multiple variables measuring donor motivations, we employ a fixed effects statistical model that guards against overfitting.[12]

We find little consistent evidence that donor motivations explain poverty selectivity generally. Many of the measures of donor motivations are associated with selectivity in directions opposite to the theoretical expectations; some do have an association in the right directions; some demonstrate no particular association with selectivity. This demonstrates an inconsistency between the set of variables ostensibly measuring the same concept, according to the existing literature. In contrast, donors' budget sizes generally albeit imperfectly exhibit relationships with poverty selectivity that are consistent with our hypothesis. Turning to the strength of the relationships assessed on the out-of-sample data, we find that donor budget sizes outperform almost all of the motivation measures in terms of out-of-sample predictive accuracy.

---

[11] Prediction or forecasting is not our goal. Rather, we are motivated to identify empirical patterns that are both consistent with theoretical explanations and generalizable to larger sets of data. We use out-of-sample fit as a tool to evaluate political processes, as advocated in recent methodological work. See Cranmer and Desmarais (2017), Hill and Jones (2014), Fariss and Jones (2017), and Schutte (2017).

[12] Specifically, we use a bootstrapped 10-fold cross-validation that includes Bayesian Model Averaging (BMA) (Montgomery & Nyhan 2010) , which we explain in the text and in the appendix.

These findings give implications for policy and study on foreign aid. The inconsistent evidence for donors' motivations implies that considerable caution is warranted when using poverty selectivity to draw inferences about how much donors care about recipient need. To be sure, we do not wish to suggest either that there is no variation in donors' motives or that motivations are irrelevant in determining poverty selectivity. After all, large bodies of work show donor motivations matter for aid allocation in general; we just find no evidence that donor motivations do so in a straightforward and consistent manner for poverty selectivity.

Rather, our findings suggest that how motivations influence poverty selectivity is more complicated than conventionally believed and that this process is likely to also involve budgetary, political considerations. Broadly speaking, this insight is in line with recent development in the literature (Bueno de Mesquita and Smith, 2009; McLean, 2015), which highlights the role of politics in translating a mix of motivations—selfish and selfless ones Heinrich (2013)—into government policy. We see it as a fruitful path to understanding donors' allocation decisions generally and poverty selectivity specifically to not focus exclusively on donors' motivations but to study how they interact with donors' characteristics, such as budget sizes, and domestic politics that underlie both donors' motivations and budgetary allocation decisions.

In the next section, we review the literature on aid allocation to develop two frameworks to understand poverty selectivity, each of which suggests a unique set variables for empirical analysis. We then subject the competing perspectives to hypothesis tests about their expected directions and evaluate the out-of-sample predictive power. Last, we conclude with a longer discussion about implications for the literatures on foreign aid and economic development.

2 Frameworks for Understanding Poverty Selectivity

Scholars and practitioners of global development have been using the concept of selectivity for quite some time to assess states' performance as aid donors. Among others, the most sustained and visible effort is the Commitment to Development Index (CDI) created by the Center for Global Development, a U.S. think tank. One indicator of the index' aid component is poverty selectivity, which simply calculates the proportion of aid going to the poorest countries. Despite its prominence, systematic attempts to examine and explain the variation in poverty selectivity have been rare. This is the task we undertake in this paper. We look to two distinct theoretical approaches in the literature on aid allocation and the broader foreign policy literature in political science. We clarify and develop each perspective to identify key determinants of poverty selectivity.

2.1 Donor Motivations Approach

The literature on foreign aid has traditionally placed a strong emphasis on donors' motivations as the central driver of aid policy. The basic idea goes back to the early studies that wanted to infer donors' motives for aid giving from patterns of aid allocation (Dudley and Montmarquette, 1976; McKinley and Little, 1979; McGillivray, 1989). The conceptual framework guiding this inferential approach is rooted in the assumption that donors vary in their motives for giving aid. Some donors are driven by development-minded motives as they want to cater to 'recipient needs.' In contrast, other donors follow self-serving motives and are concerned about commercial and security matters ('donor interests'). The fundamental assumption here is that donors' motives directly translate to different types of aid allocation.

As it stands, the framework is incomplete if the goal is to infer motivations through aid allocation patterns or, conversely, to think of motivations as a determinant of poverty selectivity. The problem is that motivations do not (yet) uniquely map to a specific aid allocation pattern. For example, suppose that, in addition to political and commercial importance of recipients, self-serving donors consider recipient poverty for their aid allocations (for non-development reasons) and give more aid to poorer recipients. Then, aid allocations favoring poor recipients become non-unique to donors with pro-development motives, rendering it impossible to infer pro-development motivations through those allocation patterns. Simultaneously, pro-development motives become a non-unique determinant of aid allocations favoring poor recipients.

Therefore, two additional assumptions are necessary for a direct, unique mapping between motivations and a specific aid allocation pattern. First, the approach must assume that political and commercial values of recipients do not influence aid allocation by a donor with a pro-development motive for some development-minded reasons. Second, it also assumes that 'recipient need' does not affect aid allocation by a donor with a non-development motive for non-development reasons.

Given that recipient poverty is one indicator of 'recipient need,' this approach would suggest that development-minded donors—and only them—favor giving more aid to poorer countries and give less aid to relatively less poor recipients. In contrast, donors with self-serving motives would be irresponsive to recipient poverty. Thus, one testable implication that follows from the *donor motivation* approach is: *donors with development-minded motivations have higher poverty selectivity.*

2.2 Foreign Policy Approach

Another set of theoretical arguments *assumes* homogenous preferences for donors and focuses on other features of donors to explain the variation in aid policy. This is in stark contrast to the motivation approach that focus exclusively on variation in donors' motivations. Rooted in the foreign policy literature (Most and Starr, 1989; Palmer and Morgan, 2011; Clark et al., 2008), these arguments treat aid strictly as a means to an end. Like other instruments of foreign policy, such as economic sanctions and military interventions, aid is regarded as the use of (scarce) governmental resources by donors to influence aspects in the international system. In this spirit, recent work by Bueno de Mesquita and Smith (2007, 2009) and Heinrich (2013) models aid as a bribe that donors pay to purchase policy concessions from other countries.

Adopting this foreign policy framework—in particular, the aid-for-policy model by Bueno de Mesquita and Smith (2007, 2009)—, we identify another donor-level determinant of poverty selectivity and derive the hypothesis associated with it. To our knowledge, the hypothesis has never been proposed in the literature. Thus, we will explain the logic behind the hypothesis.

The theory assumes that recipients' policy choices in absence of aid are optimal from the perspective of the relevant (political) decision-maker, and any shift from this optimal policy generates disutility. When a donor asks for a change in policy, the shift to a sub-optimal policy needs to be compensated via foreign aid, which the recipient government can then use to provide goods to domestic audiences in order to maintain power.[13]

One key assumption is that potential recipient countries vary in how much aid they need in exchange for making a requested policy change. One key determinant is the level of recipient wealth: in a poor country, a smaller amount of aid suffices to compensate for the dissatisfaction induced by a donor's requested policy change, *ceteris paribus*. Thus, obtaining policy changes from a poorer country is relatively cheap for a donor country whereas wealthier countries are more expensive to pay off.

The size of a donor's budget interacts with recipient wealth to determine its aid allocations. Aid allocations are funds from the donor budget that are not spent on other appreciated policies, such as health care and defense. If the government budget of the donor is larger, transferring a fixed amount away as aid is easier than if the budget were smaller. Thus, for the donor government, the costs of generating aid funds are evaluated against all other potential uses of the budget, such as the provision of health care or

[13] See Bueno de Mesquita and Smith (2009). Others make similar basic points: Girod (2012), Licht (2010), and Carter and Stone (2015).

spending on defense.[14]

Therefore, a larger budget lets a donor obtain more easily expensive policy concessions from relatively wealthier recipient countries. As a result, resource-rich donors can target wealthier recipients and tend to concentrate their aid on poorest countries less. In contrast, donors with smaller budgets can only afford to obtain cheap policy concessions from poorer recipients and therefore tend to concentrate their aid on poorer recipients more. This logic gives us another determinant of donors' poverty selectivity: *as donors' budgetary resources increase, poverty selectivity declines*.[15,16,17]

2.3 Perspectives on Poverty Selectivity

We have established two conceptual frameworks, one with a focus on donor motivations and the other on donors' resources, and have arrived at a total of nine measures that existing research expects to explain variation in poverty selectivity.

Before moving on to our empirical analysis, we restate the key difference between these conceptual frameworks. Every framework is a simplification of reality, focusing on a few essential aspects while ignoring others. The donor motivation framework focuses on donors' motivations to explain aid allocation decisions, thereby ignoring all other aspects, including political and budgetary constraints. Similarly, the foreign pol-

---

14 It merits repeating that this is neither the budget for foreign aid, nor latent resources (such as iron and steel production). The relevant budget to draw on is the amount of money available to the government to allocate to policies in a year.

15 It merits emphasizing that, according to the foreign policy approach, it is government budgets, not aid budgets, that is related to the distribution of aid across recipients. Aid budgets are endogenous to government spending and simply the sum of the aid given to all recipients.

16 The model by Bueno de Mesquita and Smith (2009), on which our framework draws from, suggests another donor-side variable besides resources to test the selectivity implications. Donors with larger winning coalitions should be less selective. Like a large budget, a large winning coalition lets the donor transfer away money more easily. However, donors in our temporal domain do not see any meaningful variation (across time and space) in the sizes of winning coalitions.

17 Our preferred operationalization would run into problems if the valuation of a recipient's concessions was to rise with the budget size, that bigger donors simply preferred to influence more with others' affairs. Some scholars argue that powerful states care more about events around the global because their power confers global interests to them. In contrast, so the argument would go, smaller states, like Iceland, do not care because of smaller ambition. However, this line of reasoning conflates power and interests, which is problematic on theoretical grounds (Palmer and Morgan, 2011). Given cheap communication technology, Iceland is surely aware about (say) trade policy elsewhere in the world and may prefer fewer trade barriers on wool, fish, and aluminum. But, as we argue, a small budget prevents it from actively meddling in it. Thus, following Palmer and Morgan (2011), we think it is analytically cleaner to not make preferences endogenous to resources, to keep preferences and saliency over policy concessions fixed in theory, and to examine the effects of varying budgetary constraints.

icy framework focuses on the role of budget constraints[18], ignoring variation in donor motivations. While both frameworks leave out essential details, it is an empirical question whether any of the determinants of selectivity that we derived from each framework is related to poverty selectivity. Our goal is to evaluate and determine which of these potential determinants of poverty selectivity captures most variation in poverty selectivity in ways that are consistent with the hypotheses while assessing their generalizability beyond particular data (Cranmer and Desmarais, 2017; Fariss and Jones, 2017).[19]

3 Research design

This research design serves two goals: testing whether the covariates' signs are in the expected direction and determining which determinants provides accurate out-of-sample predictions. As the former is routine in economics and political science, we elaborate more on the latter. Our goal is to determine how well measures of a donor's motivation and budgetary resources predict aid selectivity on observations that are not part of the data used for estimation. It is worth emphasizing that we use out-of-sample prediction with the goal to find generalizable associations (Cranmer and Desmarais, 2017; Hill and Jones, 2014; Schutte, 2017).

Accurate and powerful explanation requires robust, generalizable correlations. Our empirical approach follows in spirit the prominent efforts to examine robust, generalizable associations between predictors, which existing literature suggests, and an outcome of interest. This was first introduced by Leamer (1983, 1985) and perhaps made most famous by Sala-I-Martin (1997). In political science and economics, this approach has been applied to study, for example, civil war onset (Hegre and Sambanis, 2006), the success of economic sanctions (Bapat et al., 2013), and the decision to provide credit by the International Monetary Fund (Sturm et al., 2005). These articles' notion of robustness is captured by examining (in-sample) coefficients and p values across model specifications. We side with more recent work by Hill and Jones (2014), Schutte (2017), and Bell (2015) and examine the generalizability of associations between predictors and outcomes by relying on

[18] For the purpose of explaining poverty selectivity, we ignore other aspects of donors and donor-recipient interaction (such as their domestic institutions).

[19] There are many more determinants of aspects of donors' aid programs that the literature has examined. For example, Fuchs, Dreher and Nunnenkamp (2014)'s study examines many determinants of donor generosity that we do not include, such as the fractionalization of the donor government, the economic output gap, and colonial history, for example. While they could be associated with selectivity, they fall outside our two conceptual frameworks that guide this study. Future work could identify interesting patterns between such variables and selectivity to motivative additional theoretical research.

out-of-sample predictions.

First, we introduce the statistical model that we use to generate these predictions and evaluate the signs of covariates. Second, we discuss the operationalizations of the outcomes and predictors in great detail. Third, we present our metric for assessing generalizability.[20] Fourth, we lay out the estimation procedure.

3.1 Statistical Model

In order to generate out-of-sample predictions, we iteratively split our data into training and testing sets. Using the former, we estimate a model linking predictors to outcomes with which we calculate our measure of generalizability. This cross-validation procedure is repeated many times to average over the training-test splits. We carry out (non-parametrically) bootstrapped cross-validation to iteratively randomly split our data into subsets.

For the statistical model, we would prefer something more flexible and robust than the routinely used linear model, which is nonetheless swiftly estimable. A good candidate is Bayesian Model Averaging (BMA) (Hoeting et al., 1999; Montgomery and Nyhan, 2010). It takes predictor candidates and estimates possible sub-models based on the inclusion and exclusion of each variable. With K variables, we have $Q \equiv 2^K$ sub-models. Crucially, BMA not only considers a large space of models, it also uses Bayes' Rule to calculate the posterior probability that each potential model is the right one. Thus, it provides a principled way to attach weights to better fitting models.

In the end, the BMA prediction is simply the weighted average across the predictions from the Q models.[21] We will assess the covariates' signs by averaging the respective BMA-estimated coefficient across the ten folds. This test is conceptually and practically unrelated to any of the data that is withheld to calculate the out-of-sample variable importance, which we will introduce shortly.

3.1.1 Outcome Measures

Our outcome of interest is the poverty selectivity of a donor for a particular year. We follow the conceptual operationalization by McGillivray (1989), Roodman (2012), and

[20] The appendix offers a mathematical statement of our approach.

[21] See Hoeting et al. (1999), Montgomery and Nyhan (2010), and our appendix for more details.

Easterly and Williamson (2011):

$$y_{i,t} = 100 \cdot \frac{\sum_j \omega_{j,t} \cdot a_{c,j,t}}{\sum_j a_{i,j,t}},$$

where $y_{i,t}$ is donor i's selectivity score in year t, $a_{i,j,t}$ is donor i's aid to recipient j, and $\omega_{j,t}$ is the weight assigned to aid to recipient j. This weight is defined as a (weakly) increasing function in recipient poverty, and thus rewards aid to poorer recipients. It therefore follows that the greater the value of $y_{i,t}$, the more selective donor i's aid allocation is. A study by Clist (2015) shows that selectivity rankings can be quite sensitive to choices for $a_{i,j,t}$ and $\omega_{j,t}$. Therefore, we want to use several ways to operationalize them and then work with a factor score. Taking a factor "purges" the measurement error inherent in each of these measures.

Specifically, we use three ways to capture foreign aid ($a_{i,j,t}$): country programmable aid, aid disbursed by the donor in a year,[22] and aid committed in a particular year. For the weights ($\omega_{j,t}$) we use an indicator of whether the recipient j is among the 25 poorest countries in a year and the percentile of wealth among recipients a recipient is. We calculate both of these measures using recipient's GDP/ capita using constant dollars and using purchasing power parity (PPP).[23] Given the results by Clist (2015), a general sense of which we will echo as well, we work with the donor-year's factor score across these $3 \times 2 \times 2 = 12$ specifications for $y_{i,t}$. The appendix will show all results, some of which we will discuss in our robustness check section. Figure A.1 shows that all of the 12 measures and the factor score correlate highly (> 0.7).

We prepare these measures for the years 2000 through 2019 for 25 donors that are part of the OECD's DAC throughout those years.[24] The unit of analysis is the donor-year, and we end up with a total of 500 observations. In the appendix, we list all donors and recipients used for the calculations of the selectivity measures.[25]

---

22 This is motivated by the observations that donors not only often renege on previously committed aid (Hudson, 2013; Canavire-Bacarreza et al., 2015; Molenaers et al., 2015), but that they also disburse originally unplanned aid (when natural disasters strike). Disbursements thus also capture situational, erratic factors.

23 We thank a referee for suggesting the use of PPP.

24 Our calculation of poverty selectivity is based on 140 recipients, the usual set of developing countries whose aid receipts are tracked by the OECD DAC.

25 Across these GDP per capita observations for the recipient-years, a negligibly tiny number of them is missing, which we omit from any calculation.

3.1.2 Predictors

We require two sets of predictors (see Table 1 for a summary and data sources). The first is the budget size of the donor, and the second measures donor motivations behind aid giving. The former is operationalized by the magnitude of government spending. The size of government budget is the logarithm of government spending, which is computed by multiplying the gross domestic product and the share of the GDP that the government commands (Bueno de Mesquita and Smith, 2009).

It is not obvious which variables may measure the humanitarian concern of a donor. We turn to the voluminous literature on aid allocation to search for theoretical arguments (and not just post hoc interpretations) that particular variables capture (at least in a partial way) donors' humanitarian motives. The motivation-based approach has been extremely influential and has guided many studies to look for underlying factors forming a donor's humanitarian motives. These studies typically take a domestic actor-centric approach by focusing on some domestic actor in donor countries, which is argued to care particularly about global development, to explain the variation in some aspects of aid policy.

First, one notable example of such domestic actor-centric approach is the case of women's influence in the donor country (Breuning, 2001; Lu and Breuning, 2014; Fuchs and Richert, 2017; Kleemann et al., 2016; Hicks et al., 2016). These scholars argue that if women are more empowered in society, the country should care more about the poor abroad and provide more and higher quality of development aid. Therefore, we (first) include the share of women in the national parliament in our model. Second, this type of preferential influence may be pronounced when the minister responsible for development is female. We include a variable for this (Fuchs and Richert, 2017).

Third, another common domestic actor-centered approach focuses on political ideology of the government or the minister as a measure of a donor's humanitarian motive. Scholars argue that left-wing parties also want to help people on a global scale (Thérien and Noel, 2000; Round and Odedokun, 2004; Tingley, 2010). Survey evidence supports the observation that people holding a more left-leaning ideology tend to be more appreciative of aid (Chong and Gradstein, 2008; Paxton and Knack, 2012). Therefore, a greater influence of left-wing parties in the donor government also leads to greater commitment to development (Brech and Potrafke, 2014). Fourth, Tingley (2010) hypothesizes that right-wing parties are particularly opposed to aid, which is not a mere mirror argument. We calculate the parliamentary seat shares of left-wing and right-wing parties in the cabinet (Beck et al., 2001). Fifth, such ideology may be particularly influential when the aid-responsible minister is on the ideological left. We include a variable for this (Fuchs

Variable Names	Measurement	Source
Women Seat Share (+)	Seat share of women in the national parliament	World Development Indicators
Minister Gender (+)	Recorded as 1 if the minister reponsible for international development is female	Fuchs and Richert (2017), updated by authors
Left-Wing Seat Share (+)	Parliamentary seat share of left-wing parties in the government	Database of Political Institutions
Right-Wing Seat Share (-)	Parliamentary seat share of right-wing parties in the government	Database of Political Institutions
Minister Ideology (-)	Ideology of minister's political party on a five-point scale, going from -1 (very left) to 1 (very right)	Fuchs and Richert (2017), updated by authors
Social Spending (+)	Domestic social spending per (donor) capita	OECD
KOF Social Index (+)	Mean of de facto informational, cultural, and interpersonal dimensions of the KOF Globalization Index	Dreher (2006), Gygli et al. (2019)
Independent Agency (+)	Recorded as 1 if the aid agency works independently from the Ministry of Foreign Affairs	Fuchs et al. (2014), updated by authors
Donor Resources (-)	Logarithm of government spending (product of GDP and the share of the GDP that the government commands)	Penn World Tables

Table 1: Variable Names, Measurements, and Data Sources for the Predictors. (+)/ (-) indicate the anticipated sign on a coefficient for that variable.

and Richert, 2017).

Sixth, alongside research that focuses on political ideology of the government, there has been studies examining policy outcomes as a measure of a donor's humanitarian motive. Noël and Thérien (1995) and Thérien and Noel (2000) argue that domestic and foreign welfare and inequality considerations are driven by the same preferences. Indeed, there is evidence that domestic pro-poor policies such as increased social spending and welfare programs have a positive effect on foreign aid spending (Fuchs et al., 2014; Round and Odedokun, 2004). These policies are not necessarily embedded in ideology, may be affected by domestic institutions, and need not be on a left-right spectrum. As the home country's poverty inequality is dearer to people than the same situation abroad, we proxy the welfare emphasis by the extent of donor-side domestic social spending.

Seventh, a few scholars have also considered the extent of the donors' global, social engagement with the rest of the world (Lundsgaarde et al., 2007; Brech and Potrafke, 2014; Paxton and Knack, 2012). To measure such global connectivity at the country level, we use the average of informational, cultural and interpersonal dimensions of the KOF Globalization Index by Dreher (2006) and Gygli et al. (2019), which captures aspects of tourism, access to news, connections to foreign countries, foreign born population, as well as penetration of global brands.

Eight, while the aforementioned arguments all concern preferences of the government or the public, some studies focus on the aid agency who is typically responsible for the allocation and administration of development aid. Lancaster (2008) and Arel-Bundock et al. (2015) argue that when development aid agencies are independent from other ministries (like foreign, trade, or defense), they are less politicized and tend to be more development-oriented. We proxy this idea by using data on the *de jure* independence of the aid agency from other ministries (Fuchs and Richert, 2017).

These eight variables have been argued by different scholars to capture a donor's development motives behind foreign aid, as we reviewed above. However, merely including these measures contemporaneously may miss that their effects on selectivity accumulate over time. This has been shown in foreign aid by Thérien and Noel (2000). Therefore, we also calculate the 10-year moving average for each of these motivation variables.[26]

Generally, including all of contemporaneous and time-averaged (TA) variables in the same model would be problematic as it would raise interpretative issues and possibly bring forth multicollinearity problems. Existing research provides no theoretical guid-

[26] We do not include a moving average of donors' budget sizes. The key mechanism linking a donor's budget size to selectivity is a constraining effect. Given that such constraining can only occur contemporaneously, it would not make much theoretical sense to include averages from across ten years.

ance to favor including any over any other. Fortunately, unlike routinely used statistical models, our use of BMA is plagued less by such multicollinearity concerns. After all, it estimates sub-models so that not all variables have to be appear at the same time. Since the predictions from the models are weighted by fit, the different operationalizations will still contribute to the predictions (Montgomery and Nyhan, 2010, p. 257–261). That said, the correlation plot in Figure A.2 shows that only (some) contemporaneous and time-averaged variants of the same operationalization correlate something highly. All correlations across operationalizations are tiny.

3.1.3 Measuring Variable Importance

The explanatory power of these predictors will be judged by how much they affect out-of-sample predictions (Cranmer and Desmarais, 2017; Hill and Jones, 2014; Schutte, 2017), which measure the generalizability of the empirical association (Fariss and Jones, 2017). This is often done by calculating a metric such as the root-mean-squared error (RMSE). To assess the contribution of a *specific* variable to the predictive power of a model, a common approach is to compare the RMSE using observed data to the RMSE when the values for a single predictor are shuffled or permuted (Hill and Jones, 2014; Lu and Ishwaran, 2017). We define the *variable importance* for a variable as the change in out-of-sample RMSE associated with permuting the variable—i.e. the permuted-RMSE divided by the unpermuted RMSE.[27] Intuitively, if a variable matters, permuting the realization of the variable will result in a systematic increase in the permuted-RMSE and thus an increase in the variable importance. Conversely, if the variable matters little, it will not.

For example, suppose that we have a linear regression gives the conditional mean function of selectivity (S_j) as $E[S_j|M_j, B_j] = 1 + 0.1M_j - 5B_j$ where M_j and B_j are two predictors. Assume the observed B_j and M_j are 0.3 each, and the actual realization for the outcome is $S_j = 1$. The RMSE for this out-of-sample prediction is $\sqrt{\left(E[S_j|M_j, B_j] - S_j\right)^2} = \sqrt{(-0.47 - 1)^2} = 1.47$

Iteratively shuffling the values of variables and examining the change in out-of-sample error tells us how important a variable is for out-of-sample predictions, which gives a sense of the generalizability of the association (Cranmer and Desmarais, 2017; Fariss and Jones, 2017). Suppose we randomly permute M_j and B_j individually via a random draw from all of their respective realizations. For this illustration, suppose the permutation draw for B_j is 0.9 while retaining the actually observed value for M_j (i.e. $M_j = 0.3$). The

[27] For more detailed discussion on variable importance, see Appendix A.

new out-of-sample predictive error is 2.8, a factor ≈ 3.0 increase in predictive error (i.e. variable importance = 3.0). In contrast, if we draw the same 0.7 for M_j while using the actual B_j value, the predictive error would be 1.4. The worsening is a factor of ≈ 1.0 (i.e. no discernible change). As the predictive error increase more when B_j is permuted than M_j, the variable importance is higher for predictor B than for predictor A.[28] While it is the case that the larger variable importance coincides with the larger coefficient (and by implication t statistic), the importance estimate relies on out-of-sample data. It thus reflects the concern about the generalizability of the association.

3.2 Estimation

We have one main outcome variable and nine covariates to explain its variation. In our specification, we include the donor resources and all contemporaneous and time-average (individual) motivation-measures, ie. $1 + 2 \times 8 = 17$ included variables. Given that we are interested in explaining within-donor variation in poverty selectivity, we include a set of donor and year fixed effects in both model specifications as well as in each and every sub-models to calculate BMA predictions.

Our data contains a small amount of missingness. As has become the *de facto* gold standard for dealing with (minor) missing data (Lall, 2016), we use multiple imputation.[29] We work with 100 imputed, full data sets.

Our procedure to obtain the results are the following steps, which we repeat 1,000 times for each model specification.

1. Select an imputed data set at random.
2. Take a non-parametric bootstrap draw equal to the sample size (with replacement).
3. Randomly partition the bootstrap draw into ten folds.[30]
4. For each fold:

   (i) Estimate the model using all but the specific fold.[31]

[28] Using simulation by drawing B_j and M_j randomly from the unit-interval, the average variable importance factor is ≈ 1.6 for B_j and ≈ 1.0 for M_j.

[29] We also make use of Amelia's features to handle time-series cross-sectional data by including lags and time trends in the imputation. See our replication package for details.

[30] We enforce that the country and year fixed effects are estimable by only accepting bootstrap draws and 10-fold cross-validations that use at least five donor observations across any nine folds. In each and every sub-model used in this study, the two fixed effects were selected.

[31] We rely on the BMA package in R to estimate the model. The BMA package uses an algorithm to find the best fitting models, technically not estimating all possible sub-models. Several simulation exercises by us reveal that predictions from the BMA package are nearly identical to those when estimating all sub-models and using the Bayesian Information Criterion to calculate model weights. Simulations are

(ii) Save each averaged, mode-weighted coefficient.

(iii) Calculate the RMSE on the held-out fold.

(iv) For each predictor: permute its (and only its) values and calculate the associated RMSE.

5. Average overall and permutation-based RMSEs and coefficients across the ten estimates.
6. For each predictor: calculate variable importance by dividing the permuted-RMSE by the predictive RMSE.

4 Results

Table 2 gives a first overview of the coefficient estimates from the BMA model; we turn to the out-of-sample variable importance afterwards. The first column of estimates gives the model-averaged coefficient on the variable on the left, the range in brackets the 90% confidence interval. This coefficient is weighted-average across all sub-models explored by BMA. Since all predictors were standardize to have a mean of zero and a standard deviation of one, we can glean some strengths of associations. *Donors resources*, *KOF social index* (TA), and *social spending* (TA) have the largest coefficients. However, while the coefficient on *donor resources* is in the hypothesized direction, the other two are not. If greater *KOF social index* and *social spending* (TA, each) are indicative of more humanitarian motives, the selectivity should increase and the coefficient estimates should be positive. This is not the case for these two. This first alarms us to a pattern that we find throughout: some variables have strong associations with selectivity in a direction opposite of what the conceptual framework predicts. The last column in the table gives the share of bootstrap draws for coefficients that are in the direction suggested by the theory. Only *donor resources*, *KOF social index*, and *right-wing seat share* (both variants) have more 90% of draws in the correct direction, but all but the donor resources have small coefficient estimates.

The second column gives the model inclusion, which reveals in what share of the sub-models (averaged across bootstrap draws) the variable was included. Only *donor resources*, *KOF social index* (TA), and *social spending* (TA) are variables that were chosen generally by the BMA.[32]

Figure 1 gives a graphical depiction of the share of draws in the 'correct' direction, taken directly from the table, and a summary of the permutation-based out-of-sample

available in the replication materials. We thank a referee for making us explore this aspect of the BMA

	Coefficient	% Inclusion	% Support
Donor resources	-5.38	88.03	0.93
	[-11.54; 0.80]		
Independent agency	-0.66	30.14	0.17
	[-2.66; 0.10]		
Independent agency (TA)	0.29	18.72	0.64
	[-0.45; 1.95]		
KOF social index	2.22	41.00	0.87
	[-0.15; 6.96]		
KOF social index (TA)	-6.66	85.51	0.01
	[-12.28; -0.96]		
Left-wing seat share	-0.02	9.76	0.51
	[-0.45; 0.31]		
Left-wing seat share (TA)	-0.48	28.41	0.23
	[-2.26; 0.31]		
Minister gender	0.03	5.42	0.57
	[-0.04; 0.24]		
Minister gender (TA)	0.28	19.42	0.85
	[-0.01; 1.33]		
Minister ideology	-0.06	8.69	0.67
	[-0.43; 0.04]		
Minister ideology (TA)	0.05	9.56	0.40
	[-0.21; 0.50]		
Right-wing seat share	-0.30	23.96	0.90
	[-1.24; 0.00]		
Right-wing seat share (TA)	-1.12	58.76	0.98
	[-2.73; -0.01]		
Social spending	-0.34	13.70	0.26
	[-1.88; 0.17]		
Social spending (TA)	-4.74	85.89	0.01
	[-8.43; -0.95]		
Women seat share	-1.31	40.11	0.10
	[-4.32; 0.05]		
Women seat share (TA)	0.21	27.24	0.51
	[-2.41; 3.73]		

Table 2: Main Models of Poverty Selectivity. The sample includes 500 observations from 25 DAC donors in the period 2000 to 2019. The outcome variable is a factor based on 12 different measures of poverty selectivity (see Section 3.1.1). 'Coefficient' column gives the BMA estimates with the ranges indicating the 90% confidence intervals. Quantities summarize 1,000 BMA models based on non-parametric bootstrap draws. '% Inclusion' column represents the shares of the sub-models that included the variables. '% Support' column gives the shares of bootstrap draws for coefficients that are in the hypothesized directions. 'TA' stands for 'time averaged' (10-year average).

variable importance on the right side. Along the y-axis are the variables, and the x-axis in the left-hand panel shows the fraction of bootstrap estimates that are in the expected direction. The panel on the right-hand side gives the variable importance—i.e. the permuted out-of-sample RMSE for each variable divided by the RMSE without permutations.

package.

32 In contrast, the donor and year fixed effects for chosen in every explored sub-models across all bootstrap iterations.

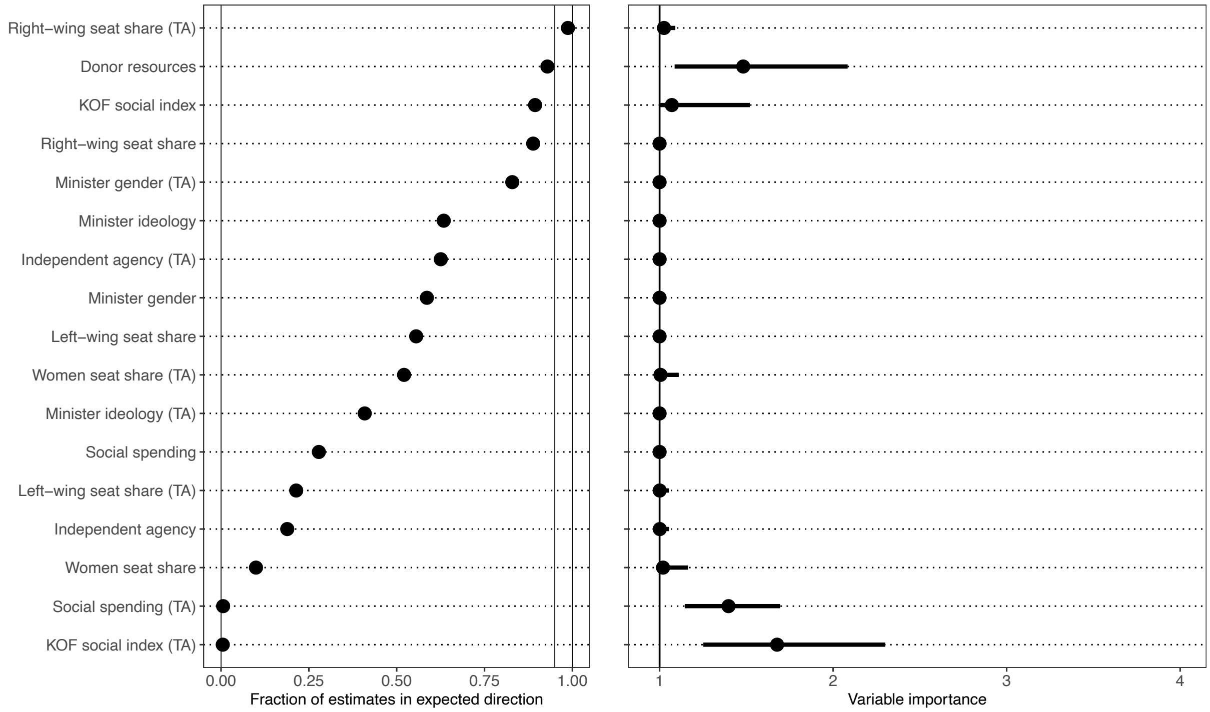

Figure 1: Percentages of Support and Variable Importance for Donor Resources and Motivation Measures The results are based on BMA estimates summarized in Table 2. The sample includes 500 observations from 25 DAC donors in the period 2000 to 2019. The outcome variable is a factor based on 12 different measures of poverty selectivity (see Section 3.1.1). In the left hand panel, the x-axis gives the the fraction of bootstrapped coefficient estimates for each of the variables on the y-axis that are in the hypothesized direction. The right hand panel shows the variable importance with 90% confidence intervals. 'TA' denotes the time-averaged calculation (10 years). The thin vertical grey line on the left hand panel indicates a fraction of 0.95.

Coefficient estimates for *donor resources* are essentially always in the hypothesized direction. The fraction of estimates that sees a negative coefficient on this variable rounds is 0.93. As a donor's budget size becomes smaller, its poverty selectivity increases and the more of its aid goes to the poorer countries. *Donor resources* also has the largest variable importance, shared with the *KOF social index* (TA). The ratio of the permuted RMSE and the unpermuted RMSE is about 1.6, meaning that the out-of-sample prediction error increases by more than half (on average) when the realizations of the donor resources are randomly shuffled. That is, the relationship between donor resources and selectivity is highly generalizable and unlikely to be a function of some unusual features of the sample.

Three motivation measures—*KOF social index*, *right-wing seat share* (both variants), *minister gender* (TA)—have high fractions of theory-consistent coefficient estimates, but their variable importance measures are small, with median essentially (or close to) zero. More troublesome is that a sizable number of coefficients have actually signs that are opposite of what the motivation approach suggests. Again, the coefficients on the motivation-variables with the largest importance scores, *social spending* (TA) and *KOF social index* (TA), are almost entirely in the unexpected directions. As we mentioned before, we have no reason to prefer any of these individual operationalizations over any other.

5 Additional results

In the appendix, we provide two additional sets of results. First, we provide results for all constituent operationalizations for our selectivity measures, which we introduced above and analyzed as a factor score. The results are shown in Section E in the appendix, where we replicate Figure 1 using each of the constituent measures. Most of the results echo the main results: *donor resources* is the variable with a consistently correctly signed coefficient and with a large VIMP. Motivation matters differ widely in the fraction of times they are signed in accordance with theory and the magnitude of the VIMP.

A noteworthy exception are some of the selectivity measures which rely on a recipient's GDP/capita measured using the Purchasing Power Parity (PPP). With the PPP-based selectivity measures, *donor resources* still tend to have the largest VIMPs with the signs of the coefficients in the hypothesized direction, but they are more heterogenous. In a few estimations, the share of (correctly) negatively signed coefficients falls to almost 25%.

However, in absence of good reasons for one aid measure (CPA, disbursement, commitment) or one way to weight recipients (percentile, bottom 25, and PPP and constant dollars), we think the most trustworthy and general estimates to us come from the using a summary measure across reasonable specifications. After all, neither theory nor substantive desiderata call specifically for one type of aid and one type of measure, which means that we should think of each of the 12 constituent measures as having measurement error of some sorts (Crabtree and Fariss, 2015). Our factor approach is a simple and common way to 'purge' such errors, obtaining what the measures have in common. By design, this comment element would be poverty selectivity.Researchers studying selectivity should pay particular attention as to whether their theory favors one measure particularly, and if not, we think a summary measure is most appropriate.

Second, our motivation estimates might suffer from post-treatment bias (King and Zeng, 2006) because donor resources might themselves be caused by (some) of the motivations measures. For example, actors that like social welfare spending might prefer a larger government as a whole, thus expanding the budget. This would generate post-treatment bias for the social welfare spending measure if the budget is also part of the regression. We re-estimate the results for the motivations variables in two ways: by dropping resources from the specification to estimate the total effect of each measure, and by relying on controlled direct effects (Acharya et al., 2016), respectively. The latter gives the effect of the motivations when (correctly) controlling for the budget channel. While some magnitudes change as we can see in Section G, our conclusions are unaffected: only a tiny number of motivation measures are consistently signed in accordance with the motivations framework and the ones signed correctly have low variable importance. Some motivation measures are signed consistently opposite to what the motivations framework would suggest, at times with large variable importance measures.

6 Discussion

Our analysis shows that donor resources contribute explanatory power and exhibit a relationship with poverty selectivity that generalizes beyond particular data. This result suggests, in line with recent work on foreign aid and general theories of foreign policy (Bueno de Mesquita and Smith, 2009; Palmer and Morgan, 2011), that we should acknowledge not only the crucial roles of budgetary and political constraints in foreign aid decisions, but also in poverty selectivity more specifically.

Our results also show that direct evidence for an influential conceptual framework in the literature on foreign aid is strikingly weak at least in the context of poverty selectivity. This opens up a vexing question: how could a donor's motivations fail to associate with selectivity despite the plausible intuition that they ought to? One possibility is that our indicators are just poor operationalizations. While that is possible, we believe this is unlikely. After all, we searched the vast existing literature and used all donor-level predictors that scholars have argued to affect genuine development-mindedness. Either the aid literature has overlooked for decades a powerful driver behind motivations, or motivations matter little for selectivity.

Alternatively, we believe that the absence of a consistently close correspondence between motivations and selectivity likely stems from what the *motivations* approach tends to ignore, namely a notion of politics and budgetary considerations. This is consistent with the recent developments in the literature on foreign aid that highlights how consequential both omissions ought to be (Bueno de Mesquita and Smith, 2009; Wright and Winters, 2010; McLean, 2015; Milner and Tingley, 2010; Heinrich, 2013; Carter and Stone, 2015; Heinrich et al., 2018).

These null results are of practical importance as well. Our findings suggest that a note of caution is in order while drawing inferences about donor motives based on poverty selectivity. Interested in learning about motivations behind foreign aid, scholars have used regressions to model donors' aid allocations as a function of 'recipient need' (e.g. GDP per capita) and interpret the resulting regression coefficient estimates to mean how much donors care about recipient need. However, using selectivity to measure donor motives can be misleading and may represent something different, according to our evidence.

While selectivity appears barely associated with motivations behind, one should not lose sight that motivations are related to the effectiveness of aid (Bermeo, 2011; Bearce and Tirone, 2010). Since it is effectiveness and selectivity in itself that matters more for global development, our results do not suggest tanking the emphasis on donors' motivations but rather considering them in the context of models of politics. Political science offers many ways to measure motivations besides selectivity as reviewed above. Scholars who have connected donor-side motivations to aid outcomes theorize that these measures should captures motives, often taking the laudable step to bolster arguments by drawing on surveys and other research designs that can capture motivation (Fuchs and Richert, 2017). The marrying of political considerations, such as public opinion, attention, and budget constraints, and motivations has led to fruitful insight in foreign aid (Eisensee and Strömberg, 2007; Heinrich et al., 2018). We view this as a more fruitful way forward in examining the how motives condition effectiveness of aid.

7 Conclusion

It is believed that donors could make more strides toward ending extreme poverty by reallocating aid to countries with greatest poverty. Major global anti-poverty initiatives are built on this premise. To achieve these goals, activists and scholars criticize donors for their failure to target their aid to the poorest countries and for a lack of commitment to the goal of eradicating poverty. That is, they primarily aim to improve donor's development motivation.

While such a motivation-centric approach is intuitive and prevalent, our findings suggest that other considerations, namely budget constraints and recipient-donor bargaining, may play crucial roles in how the aid policy targets the poor. This implies that if the goal is to improve donors' commitment to global development and the effectiveness of aid, activists ought not to focus on selectivity as a metric. Rather, activists may do better by relying on more direct measures of motivations as performance metrics and focusing on improving factors that are more closely connected to measures of motivations (Bearce and Tirone, 2010; Dreher et al., 2014; Dunning, 2004).

We do not wish to suggest that donor motivations are irrelevant in determining how donors allocate aid. Indeed, we know there is variation in donor motivations and we have strong theoretical reasons to expect that they matter in aid allocation. What our findings show is that how donor motivations manifest themselves in poverty selectivity is not as straightforward as our existing conceptual framework suggests. In line with some new developments in the literature (Bueno de Mesquita and Smith, 2009; Heinrich, 2013; McLean, 2015), we believe that future research would benefit from considering both donor motivations and budgetary and political considerations under a single framework.

References

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3):512–529.

Alesina, A. and Dollar, D. (2000). Who gives foreign aid to whom and why? *Journal of Economic Growth*, 5(1):33–63.

Arel-Bundock, V., Atkinson, J., and Potter, R. A. (2015). The limits of foreign aid diplomacy: How bureaucratic design shapes aid distribution. *International Studies Quarterly*, 59(3):544–556.

Bapat, N. A., Heinrich, T., Kobayashi, Y., and Morgan, T. C. (2013). Determinants of sanctions effectiveness: Sensitivity analysis using new data. *International Interactions*, 39(1):79–98.

Bearce, D. H. and Tirone, D. C. (2010). Foreign aid effectiveness and the strategic goals of donor governments. *The Journal of Politics*, 72(3):837–851.

Beck, T., Clarke, G., Groff, A., Keefer, P., and Walsh, P. (2001). New tools in comparative political economy: The database of political institutions. *The World Bank Economic Review*, 15(1):165–176.

Bell, M. S. (2015). Examining explanations for nuclear proliferation. *International Studies Quarterly*, 60(3):520–529.

Bermeo, S. B. (2011). Foreign aid and regime change: A role for donor intent. *World Development*, 39(11):2021–2031.

Bermeo, S. B. (2017). Aid allocation and targeted development in an increasingly connected world. *International Organization*, pages 1–32.

Brech, V. and Potrafke, N. (2014). Donor ideology and types of foreign aid. *Journal of Comparative Economics*, 42(1):61–75.

Breuning, M. (2001). Women's representation and development assistance: a cross-national study. *Women & Politics*, 23(3):35–55.

Briggs, R. C. (2017). Does foreign aid target the poorest? *International Organization*, 71(1):187–206.

Briggs, R. C. (2018). Poor targeting: A gridded spatial analysis of the degree to which aid reaches the poor in africa. *World Development*, 103(3):133–148.

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

Bueno de Mesquita, B. and Smith, A. (2009). A political economy of aid. *International Organization*, 63(2):309–340.

Canavire-Bacarreza, G. J., Neumayer, E., and Nunnenkamp, P. (2015). Why aid is unpredictable: An empirical analysis of the gap between actual and planned aid flows. *Journal of International Development*, 27(4):440–463.

Carter, D. B. and Stone, R. W. (2015). Democracy and multilateralism: The case of vote buying in the un general assembly. *International Organization*, 69(1):1–33.

Chong, A. and Gradstein, M. (2008). What determines foreign aid? the donors' perspective. *Journal of Development Economics*, 87(1):1–13.

Clark, D. H., Nordstrom, T., and Reed, W. (2008). Substitution is in the variance: Resources and foreign policy choice. *American Journal of Political Science*, 52(4):763–773.

Clist, P. (2011). 25 years of aid allocation practice: Whither selectivity? *World Development*, 39(10):1724–1734.

Clist, P. (2015). Do performance measures of donors' aid allocation underperform? *The World Economy*, 38(5):805–824.

Collier, P. and Dollar, D. (2002). Aid allocation and poverty reduction. *European Economic Review*, 46(8):1475–1500.

Crabtree, C. D. and Fariss, C. J. (2015). Uncovering patterns among latent variables: human rights and de facto judicial independence. *Research & Politics*, 2(3):2053168015605343.

Cranmer, S. J. and Desmarais, B. A. (2017). What can we learn from predictive modeling? *Political Analysis*, 25(2):145–166.

Dollar, D. and Levin, V. (2006). The increasing selectivity of foreign aid, 1984–2003. *World development*, 34(12):2034–2046.

Dreher, A. (2006). Does globalization affect growth? evidence from a new index of globalization. *Applied Economics*, 38(10):1091–1110.

Dreher, A., Eichenauer, V. Z., and Gehring, K. (2014). Geopolitics, aid and growth.

Dudley, L. and Montmarquette, C. (1976). A model of the supply of bilateral foreign aid. *The American Economic Review*, 66(1):132–142.

Dunning, T. (2004). Conditioning the effects of aid: Cold war politics, donor credibility, and democracy in africa. *International Organization*, 58(2):409–423.

Easterly, W. (2007). Are aid agencies improving? *Economic Policy*, 22(52):634–678.

Easterly, W. and Williamson, C. R. (2011). Rhetoric versus reality: the best and worst of aid agency practices. *World Development*, 39(11):1930–1949.

Eisensee, T. and Strömberg, D. (2007). News droughts, news floods, and us disaster relief. *The Quarterly Journal of Economics*, 122(2):693–728.

Fariss, C. J. and Jones, Z. M. (2017). Enhancing validity in observational settings when replication is not possible. *Political Science Research and Methods*, pages 1–16.

Finnemore, M. (1997). Redefining development at the world bank. *International development and the social sciences: Essays on the history and politics of knowledge*, pages 203–27.

Fuchs, A., Dreher, A., and Nunnenkamp, P. (2014). Determinants of donor generosity: A survey of the aid budget literature. *World Development*, 56:172–199.

Fuchs, A. and Richert, K. (2017). Development minister characteristics and aid giving. *European Journal of Political Economy*.

Girod, D. M. (2012). Effective foreign aid following civil war: The nonstrategic-desperation hypothesis. *American Journal of Political Science*, 56(1):188–201.

Gygli, S., Haelg, F., Potrafke, N., and Sturm, J.-E. (2019). The kof globalisation index–revisited. *The Review of International Organizations*, 14(3):543–574.

Hegre, H. and Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution*, 50(4):508–535.

Heinrich, T. (2013). When is foreign aid selfish, when is it selfless? *The Journal of Politics*, 75(2):422–435.

Heinrich, T., Kobayashi, Y., and Long, L. (2018). Voters get what they want (when they pay attention): Human rights, policy spoils, and foreign aid. *International Studies Quarterly*, 62(1):195–207.

Hicks, D. L., Hicks, J. H., and Maldonado, B. (2016). Women as policy makers and donors: Female legislators and foreign aid. *European Journal of Political Economy*, 41:46–60.

Hill, D. W. and Jones, Z. M. (2014). An empirical evaluation of explanations for state repression. *American Political Science Review*, 108(3):661–687.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical Science*, 14(4):382–401.

Honig, D. and Weaver, C. (2019). A race to the top? the aid transparency index and the social power of global performance indicators. *International Organization*, 73(3):579–610.

Hudson, J. (2013). Promises kept, promises broken? the relationship between aid commitments and disbursements. *Review of Development Finance*, 3(3):109–120.

Kelley, J. G. and Simmons, B. A. (2015). Politics by number: indicators as social pressure in international relations. *American Journal of Political Science*, 59(1):55–70.

King, G. and Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, 14(2):131–159.

Kleemann, L., Nunnenkamp, P., and Thiele, R. (2016). Gender inequality, female leadership and aid allocation. *Journal of International Development*, 28(3):376–395.

Knack, S., Rogers, F. H., and Eubank, N. (2011). Aid quality and donor rankings. *World Development*, 39(11):1907–1917.

Lall, R. (2016). How multiple imputation makes a difference. *Political Analysis*, 24(4):414–433.

Lancaster, C. (2008). *Foreign aid: Diplomacy, development, domestic politics*. University of Chicago Press.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73(1):31–43.

Leamer, E. E. (1985). Sensitivity analyses would help. *The American Economic Review*, 75(3):308–313.

Licht, A. A. (2010). Coming into money: The impact of foreign aid on leader survival. *Journal of Conflict Resolution*, 54(1):58–87.

Lu, K. and Breuning, M. (2014). Gender and generosity: does women's representation affect development cooperation? *Politics, Groups, and Identities*, 2(3):313–330.

Lu, M. and Ishwaran, H. (2017). A machine learning alternative to p-values. *arXiv preprint arXiv:1701.04944*.

Lundsgaarde, E., Breunig, C., and Prakash, A. (2007). Trade versus aid: donor generosity in an era of globalization. *Policy sciences*, 40(2):157–179.

Maizels, A. and Nissanke, M. K. (1984). Motivations for aid to developing countries. *World Development*, 12(9):879–900.

McGillivray, M. (1989). The allocation of aid among developing countries: A multi-donor analysis using a per capita aid index. *World Development*, 17(4):561–568.

McKinley, R. D. and Little, R. (1979). The us aid relationship: a test of the recipient need and the donor interest models. *Political Studies*, 27(2):236–250.

McLean, E. V. (2015). Multilateral aid and domestic economic interests. *International Organization*, 69(1):97–130.

Milner, H. V. and Tingley, D. H. (2010). The political economy of us foreign aid: American legislators and the domestic politics of aid. *Economics & Politics*, 22(2):200–232.

Molenaers, N., Gagiano, A., Smets, L., and Dellepiane, S. (2015). What determines the suspension of budget support? *World Development*, 75:62–73.

Montgomery, J. M. and Nyhan, B. (2010). Bayesian model averaging: Theoretical developments and practical applications. *Political Analysis*, 18(2):245–270.

Most, B. and Starr, H. (1989). *Inquiry, Logic, and International Politics*. University of South Carolina.

Murdie, A. M. and Davis, D. R. (2012). Shaming and blaming: Using events data to assess the impact of human rights ingos. *International Studies Quarterly*, 56(1):1–16.

Noël, A. and Thérien, J.-P. (1995). From domestic to international justice: the welfare state and foreign aid. *International Organization*, 49(03):523–553.

Nunnenkamp, P. and Thiele, R. (2006). Targeting aid to the needy and deserving: nothing but promises? *The World Economy*, 29(9):1177–1201.

Palmer, G. and Morgan, T. C. (2011). *A theory of foreign policy*. Princeton University Press.

Paxton, P. and Knack, S. (2012). Individual and country-level factors affecting support for foreign aid. *International Political Science Review*, 33(2):171–192.

Riddell, R. C. (2007). *Does foreign aid really work?* Oxford University Press.

Roodman, D. (2012). An index of donor performance. Center for Global Development Working Paper 67.

Round, J. I. and Odedokun, M. (2004). Aid effort and its determinants. *International Review of Economics & Finance*, 13(3):293–309.

Sala-I-Martin, X. X. (1997). I just ran two million regressions. *The American Economic Review*, 87(2):178–183.

Schutte, S. (2017). Regions at risk: predicting conflict zones in african insurgencies. *Political Science Research and Methods*, 5(3):447–465.

Sturm, J.-E., Berger, H., and De Haan, J. (2005). Which variables explain decisions on imf credit? an extreme bounds analysis. *Economics & Politics*, 17(2):177–213.

Thérien, J.-P. and Noel, A. (2000). Political parties and foreign aid. *American Political Science Review*, 94(1):151–162.

Tingley, D. (2010). Donors and domestic politics: Political influences on foreign aid effort. *The quarterly review of economics and finance*, 50(1):40–49.

Wang, Y. (2016). The effect of bargaining on us economic aid. *International Interactions*, 42(3):479–502.

Wright, J. and Winters, M. (2010). The politics of effective foreign aid. *Annual Review of Political Science*, 13:61–80.

Evaluating Explanations for Poverty Selectivity in Foreign Aid

Appendix

A Estimation procedure

In this section, we lay out the details for the estimation procedure and how we calculate the variable importance measures reported in the paper.

Let N^f be the set of all observations in the data, with each observation indexed by i. Taking a non-parametric bootstrap draw gives a novel set of indices, N^{bs}. When we randomly split N^{bs} into ten folds, we obtain ten subsets of N^{bs} denoted by $N_1^{bs}, ..., N_{10}^{bs}$.

Let $m \in \{1, ..., 10\}$ be the index for the partition. The $N^{bs} \setminus N_m^{bs}$ observations are used to estimate a statistical model, which gives a conditional mean function for the outcome y_i,

$$E(y_i|x_{i,k}, \mathbf{x}_{i,-k}) = g\left(x_{i,k}, \mathbf{x}_{i,-k}\right).$$

This predictive function relates the pair predictor $(x_{i,k}, \mathbf{x}_{i,-k})$ to the outcome. In our case, the outcome is a measure of how selective a donor is, and the predictors contain operationalizations of donor intent and budget size (plus other covariates).

The statistical model that we employ is Bayesian Model Averaging (BMA).[33] It estimates all submodels that are possible given the inclusion and exclusion of the K predictors. A generic submodel q is

$$\mathcal{M}_q : y_i = \beta_{0,q} + \sum_k x_{i,k}\beta_{k,q} + \epsilon_{i,q}.$$

Some elements in $\beta_{k,q}$ are restricted to be zero which is equivalent to dropping the corresponding predictor. Thus, BMA considers $Q \equiv 2^K$ models.

A crucial feature of BMA is that it upweights outputs from models that fit better. With a prior belief of $\pi(\mathcal{M}_q)$ that the q^{th} model is the correct model, the posterior belief w_q is:

$$w_q = p(\mathcal{M}_q|Y) = \frac{p(Y|\mathcal{M}_q)\pi(\mathcal{M}_q)}{\sum_q p(Y|\mathcal{M}_q)\pi(\mathcal{M}_q)},$$

with Y the vector of $y_1, ..., y_N$. A higher posterior belief suggests we should upweight the results from this model. In practice (and for the sake of speed), we use an approximation for these weights by using the Bayesian Information Criterion (BIC) for each $\mathcal{M}_q$.

Putting each model's estimates and the models' posterior probabilities together, the conditional mean function for an estimated BMA model is

$$g\left(x_{i,k}, \mathbf{x}_{i,-k}\right) = \sum_q \left(w_q \left(\beta_{q,0} + x_{i,k}\beta_{k,q} + x_{i,-k}\beta_{q,-k}\right)\right).$$

33 See Montgomery and Nyhan (2010) for an introduction in political science to BMA.

Averaging across the observations that were withheld from estimation N_m^{bs} gives the out-of-sample root-mean-square error for the m^{th} fold:[34]

$$\Delta_0 = \sqrt{Avg_{i \in N_m^{bs}} \left(y_i - g\left(x_{i,k},\ \mathbf{x}_{i,-k}\right)\right)^2}.$$

For each of the K variables, we permute that variable's realizations. Let p be a vector of length $|N^{bs}|$ with each element being a randomly drawn index. Using the withheld fold's observations (with the values of the k^{th} variable permuted), we calculate a new out-of-sample root-mean-square error:

$$\Delta_k = \sqrt{Avg_{i \in N_m} \left(y_i - g\left(x_{p_i,k},\ \mathbf{x}_{i,-k}\right)\right)^2}.$$

This lets us define the variable importance index for variable k in the estimate:

$$\mathbf{VIMP}_k = \Delta_k / \Delta_0,$$

which we average across the ten-folds. The division normalizes the scale which matters because we use several outcome measures. The measure captures: *relative* to the best out-of-sample prediction, by how much does noising up x_k worsen the prediction? For $\mathbf{VIMP}_k$ values at around one, the out-of-sample error is largely unaffected by permuting the variable k, suggesting that it matters little to capture variation in the outcome. In contrast, a large $\mathbf{VIMP}_k$ tells us that x_k is responsible for more variation in the outcome. Scholars seeking to understand the outcome phenomenon should focus on variables with higher $\mathbf{VIMP}_k$.

[34] Avg_b is calculating the average over the indices b.

B Donors and recipients in the data

Donor countries

Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, South Korea, Spain, Sweden, Switzerland, United Kingdom, and United States

Recipient countries

Cuba, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago, Barbados, Dominica, Grenada, Saint Lucia, Saint Vincent and the Grenadines, Antigua and Barbuda, Saint Kitts and Nevis, Mexico, Belize, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Guyana, Suriname, Ecuador, Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Malta, Albania, Montenegro, Macedonia, Croatia, Slovenia, Moldova, Ukraine, Belarus, Armenia, Georgia, Azerbaijan, Cabo Verde, Sao Tome and Principe, Guinea-Bissau, Equatorial Guinea, Gambia, Mali, Senegal, Benin, Mauritania, Niger, Cote d'Ivoire, Guinea, Burkina Faso, Liberia, Sierra Leone, Ghana, Togo, Cameroon, Nigeria, Gabon, Central African Republic, Chad, Congo, Uganda, Kenya, Tanzania, Burundi, Rwanda, Djibouti, Ethiopia, Eritrea, Angola, Mozambique, Zambia, Zimbabwe, Malawi, South Africa, Namibia, Lesotho, Botswana, Swaziland, Madagascar, Comoros, Mauritius, Seychelles, Morocco, Algeria, Tunisia, Libya, Sudan, Iran, Turkey, Iraq, Egypt, Lebanon, Jordan, Saudi Arabia, Yemen, Bahrain, Oman, Afghanistan, Turkmenistan, Tajikistan, Kyrgyzstan, Uzbekistan, Kazakhstan, China, Mongolia, Korea, India, Bhutan, Pakistan, Bangladesh, Sri Lanka, Maldives, Nepal, Thailand, Cambodia, Laos, Vietnam, Malaysia, Philippines, Indonesia, Timor-Leste, Papua New Guinea, Vanuatu, Solomon Islands, Kiribati, Tuvalu, Fiji, Tonga, Marshall Islands, Palau, Micronesia, and Samoa.

C Descriptive statistics

C.1 Correlations between outcomes

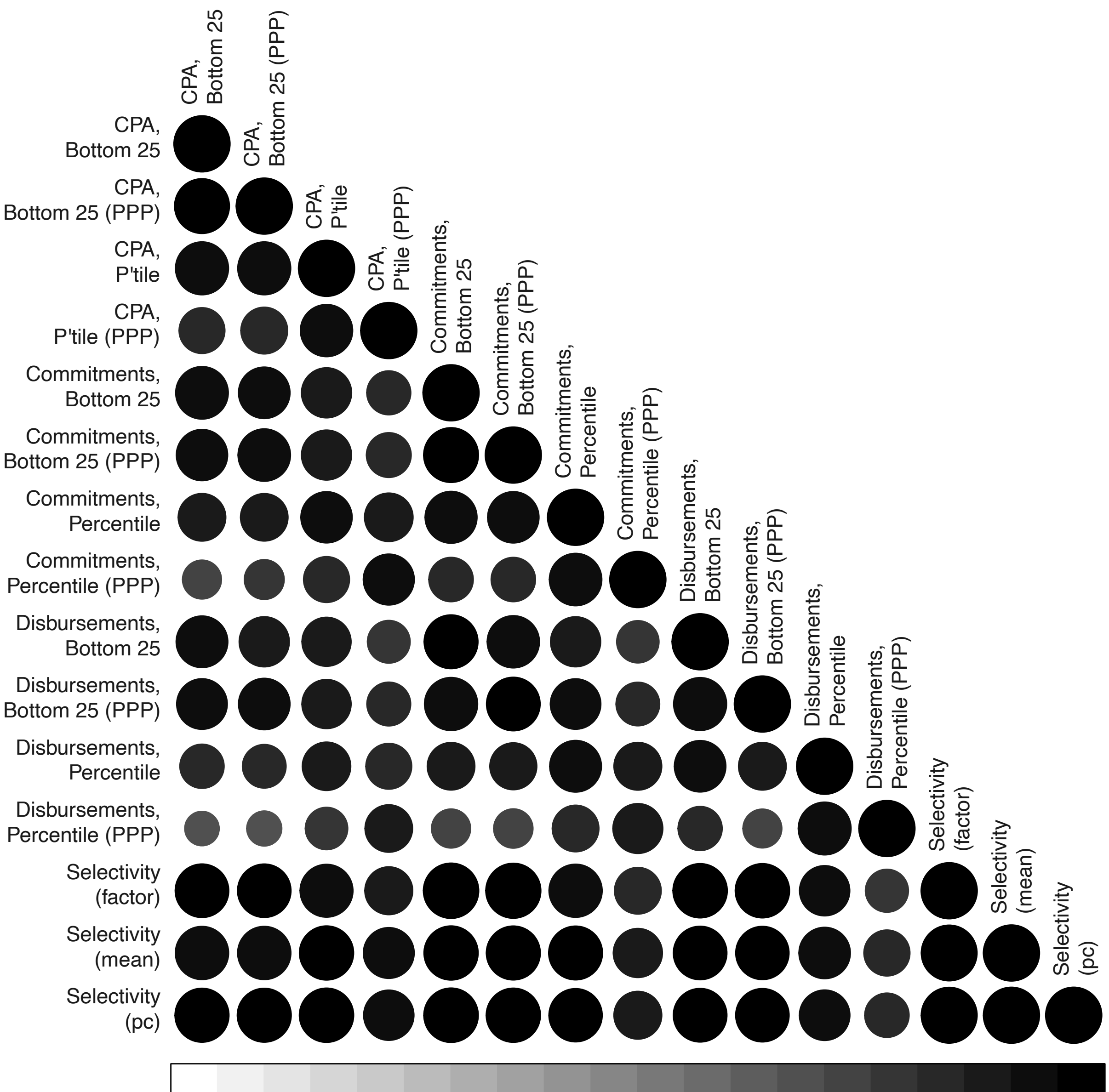

Figure A.1: Correlations between outcome measures. The color and size corresponds to the correlation. See the legend at the bottom. Estimates averaged over all imputed data sets.

C.2 Correlations between all predictors

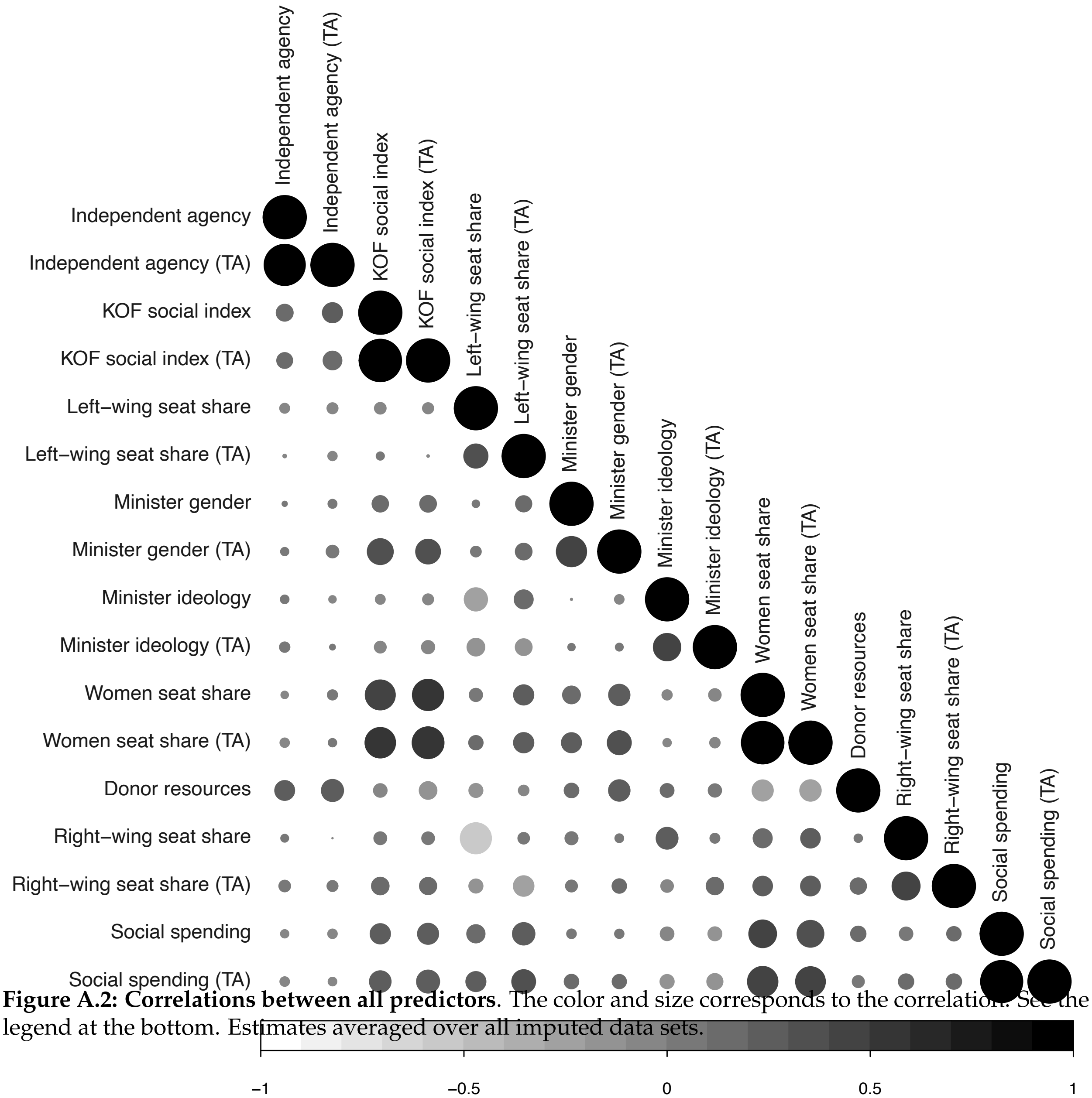

Figure A.2: Correlations between all predictors. The color and size corresponds to the correlation. See the legend at the bottom. Estimates averaged over all imputed data sets.

D Out-of-sample RMSE by model

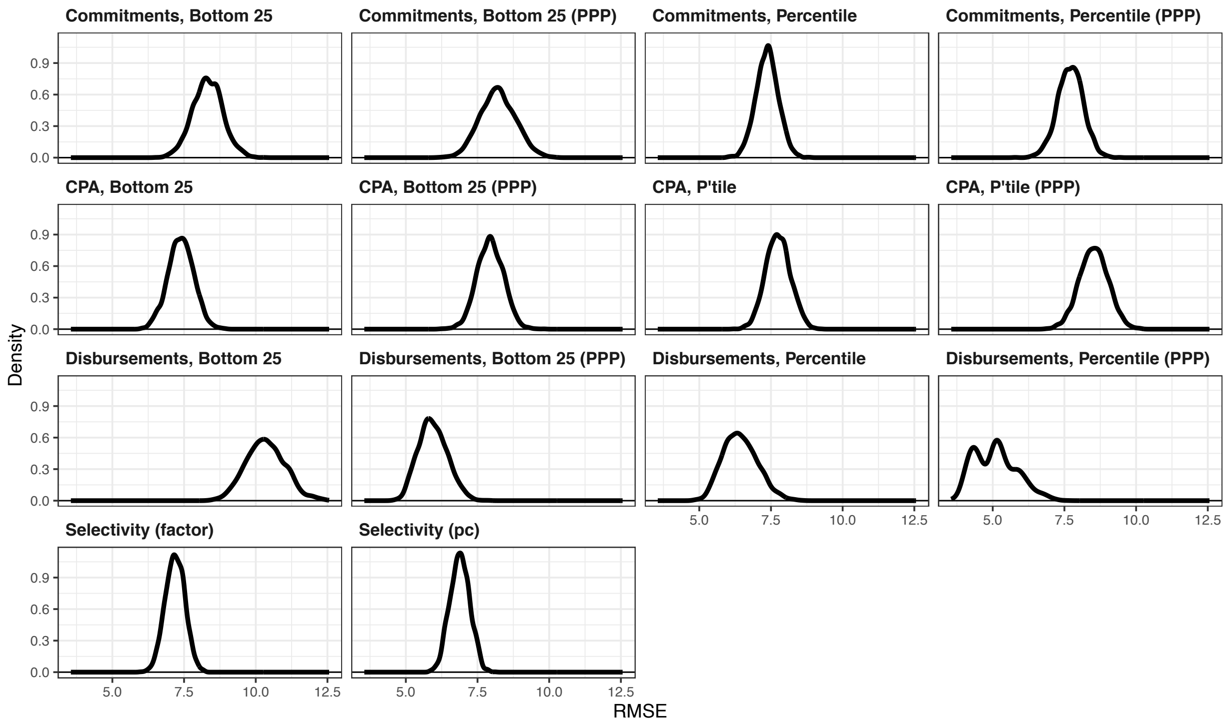

Figure A.3: Out-of-sample RMSE for each model. Each panel gives the out-of-sample RMSE error of a model using the outcome stated atop. Uncertainty comes from the non-parametric bootstrap draws.

E Additional Results I: All Selectivity Measures

Principal Component of Selectivity Results

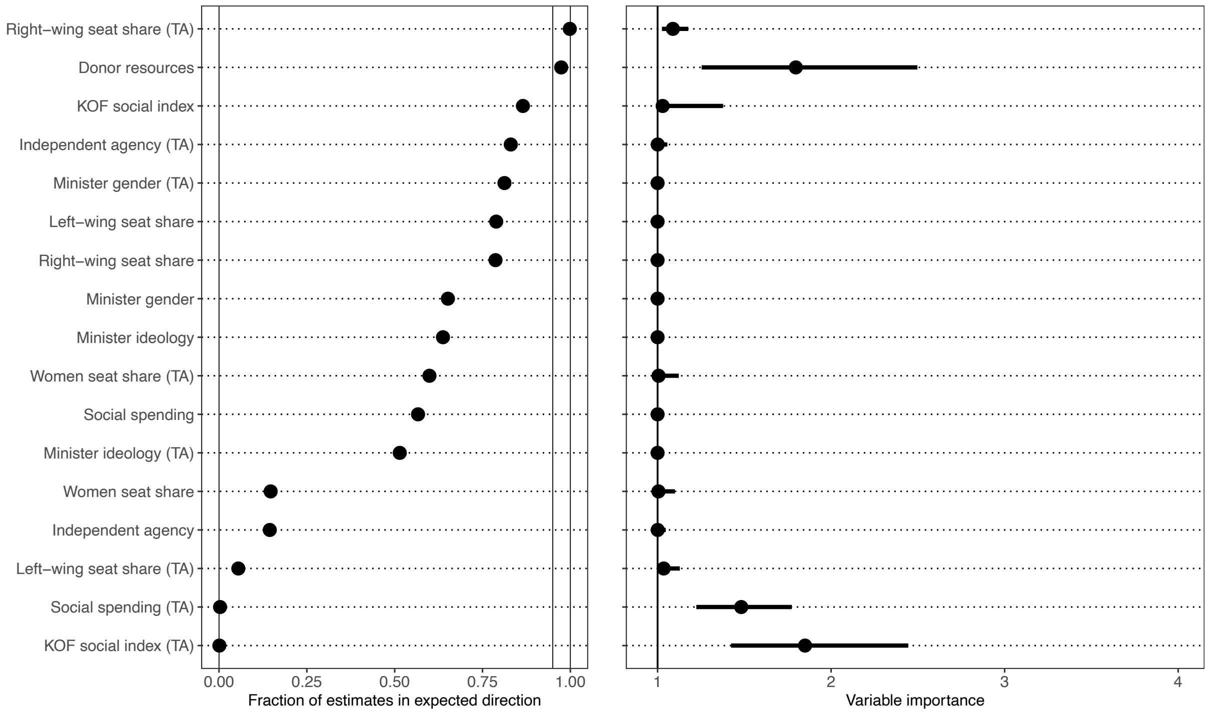

Figure A.4: Results for selectivity measured via the principal component of basic measures. The figure gives the fraction of non-parametric bootstrap estimates of the variables' signs that are in the direction consistent with the theoretical framework (left side) and the corresponding variable importance (right side). The gives the 90% confidence interval.

Aid Commitments Selectivity Results

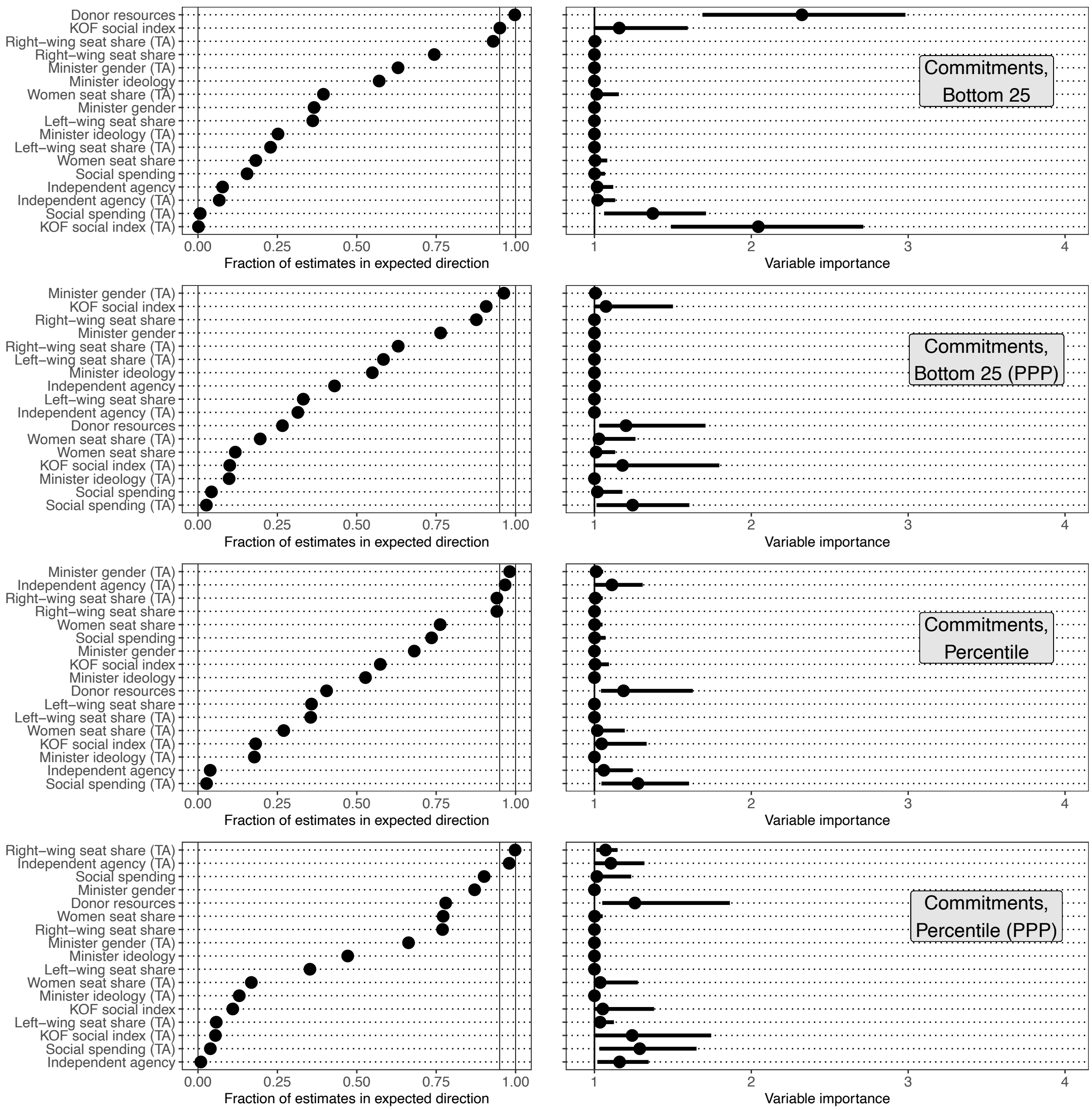

Figure A.5: Results for selectivity measured via aid commitments. Each panel in the figure gives the fraction of non-parametric bootstrap estimates of the variables' signs that are in the direction consistent with the theoretical framework (left side) and the corresponding variable importance (right side). The gives the 90% confidence interval. Each of panels uses a different metric to measure recipients' wealth ($\omega_{j,t}$).

Aid Disbursement Selectivity Results

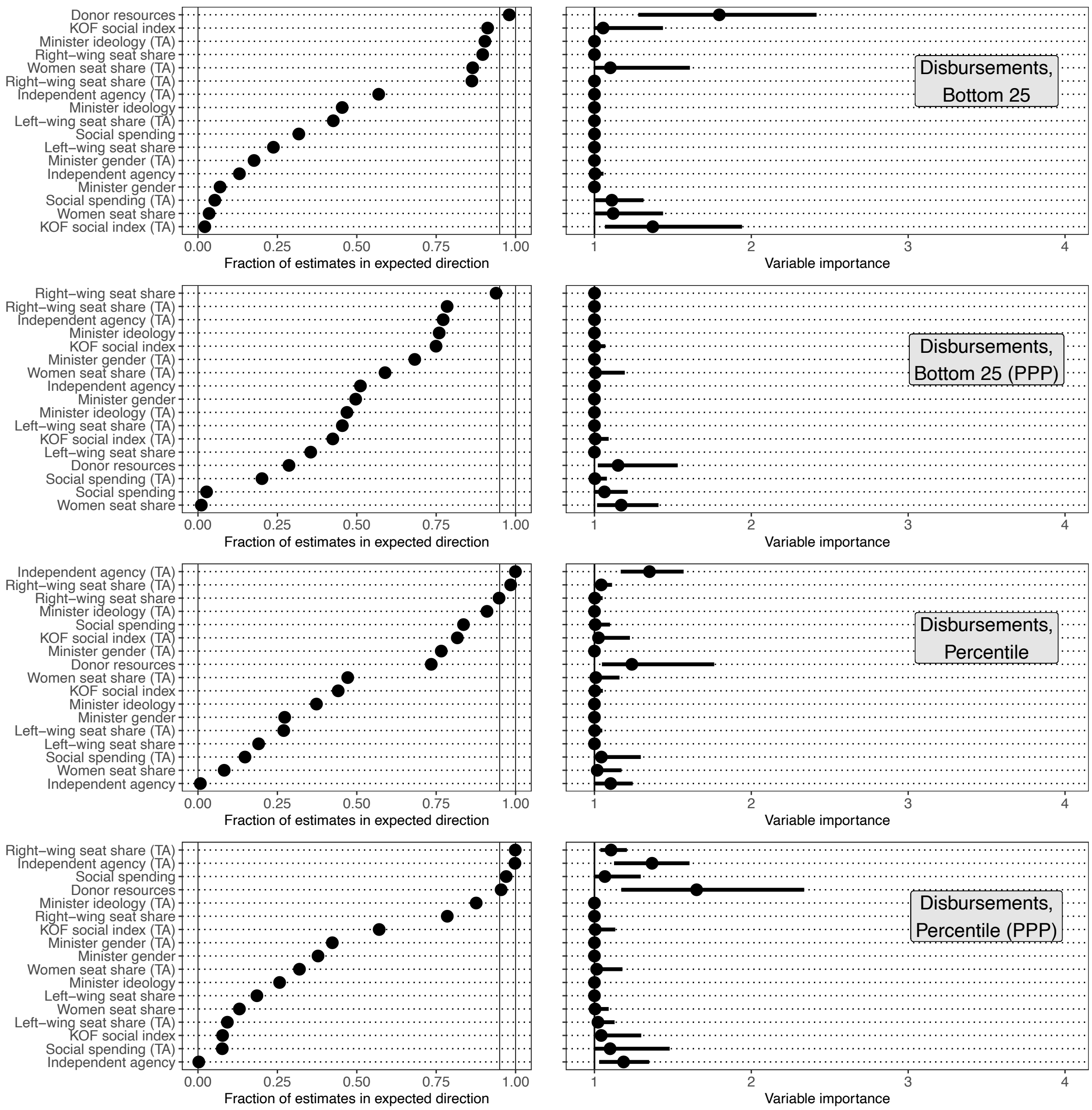

Figure A.6: Results for selectivity measured via aid disbursements. Each panel in the figure gives the fraction of non-parametric bootstrap estimates of the variables' signs that are in the direction consistent with the theoretical framework (left side) and the corresponding variable importance (right side). The gives the 90% confidence interval. Each of panels uses a different metric to measure recipients' wealth ($\omega_{j,t}$).

Country Programmable Aid Selectivity Results

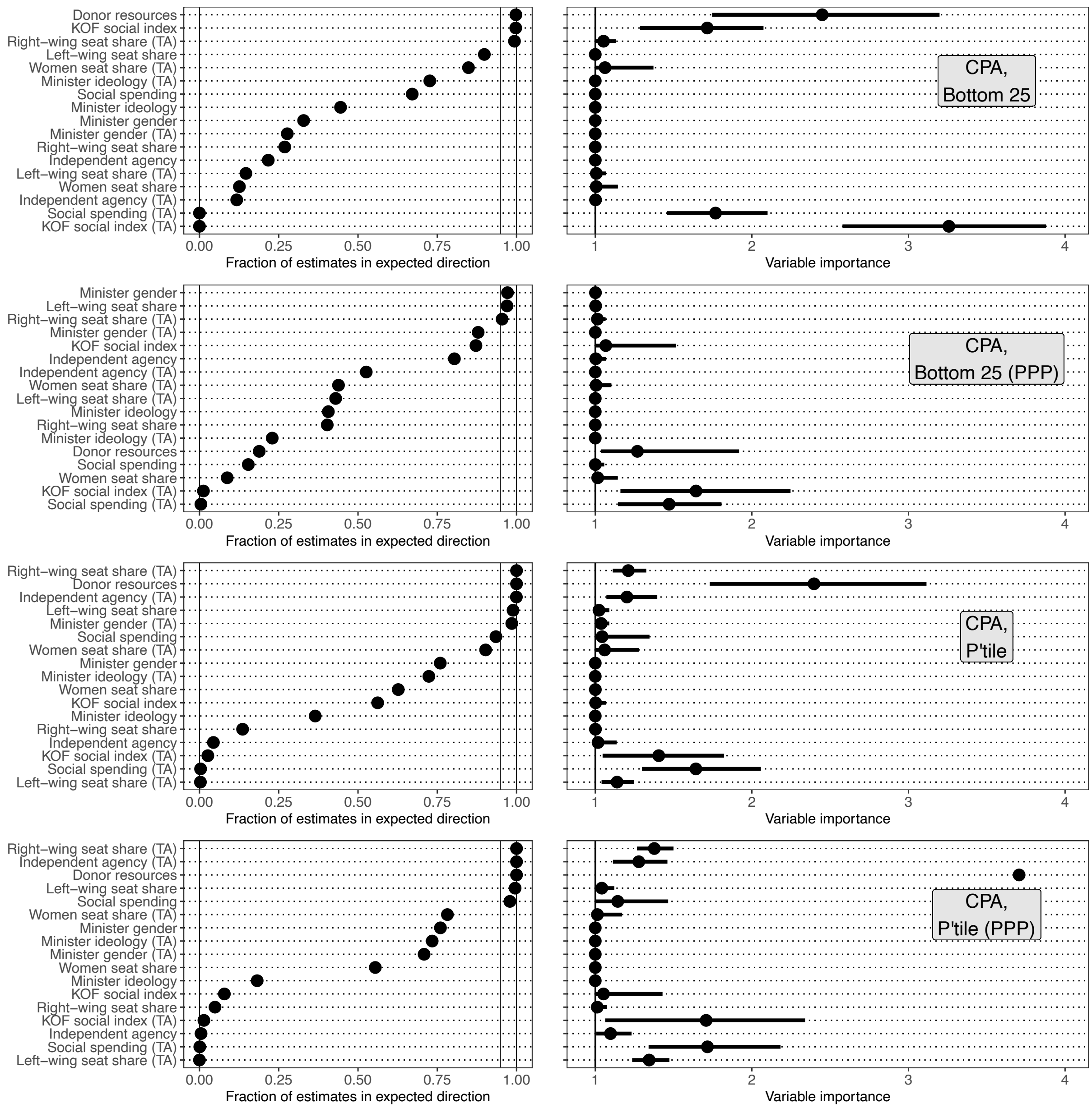

Figure A.7: Results for selectivity measured via country programmable aid. Each panel in the figure gives the fraction of non-parametric bootstrap estimates of the variables' signs that are in the direction consistent with the theoretical framework (left side) and the corresponding variable importance (right side). The gives the 90% confidence interval. Each of panels uses a different metric to measure recipients' wealth ($\omega_{j,t}$).

F Additional Results II: Ranking the Variable Importance Results

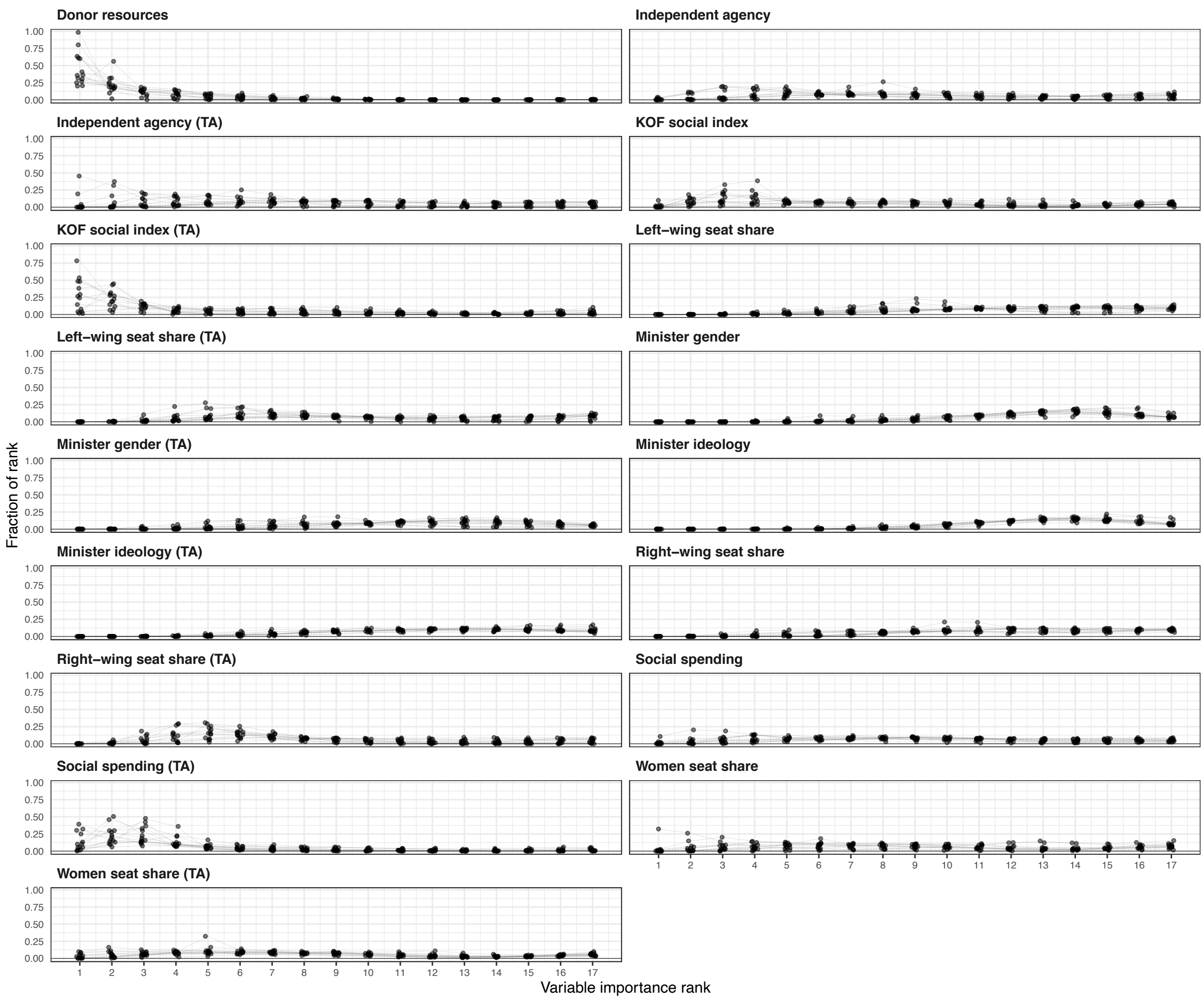

Figure A.8: Ranks of variable importance across all selectivity outcomes. Each panel shows the fraction of bootstrap estimates for variable importance that occupy each rank (listed along the x-axis) for a predictor. Each model's estimates are shown in a thin, dotted lines.

G Additional Results III: Alternative estimates for donor motivations

There is the concern that some of the measures of donor motivations might themselves be causes of the size of the government budget. For example, a greater welfare state or a long history of left-wing parties in power may leave the budget enlarged. If this is the case, part of the effects of the donor motivations variables are currently being picked by the donor budget variable. That is, the estimates for (some of) the donor motivations variables above may be suffering from post-treatment bias in unknowable directions.[35]

We tackle this issue in two ways. First, we re-estimate the factor-based selectivity model with all the motivations measures while dropping the donor resources. This approach would give us the "total effects" of the motivations variables. Second, while the inclusion of the donor budget variable may induce post-treatment bias, omission of the donor resources from the regression may generate an omitted variable problem. This tricky twin problem is not solvable in standard linear models, including our BMA approach. Thus, we address this dual conundrum by estimating controlled direct effects of motivations variables following Acharya et al. (2016). In a two-step procedure called the sequential g-estimator, we first estimate the marginal effect of donor budgets on the selectivity outcome and then "purges" the selectivity outcome of this effect (i.e. "holding donor budget constant"). Next, we regress this outcome free of the effects of donor resources on the motivation variables. This gives the direct effects of the motivations variables when "controlling" for the effect of the donor resources.

By and large, little changes for the variable importance and fraction of estimates in the correct direction under either approach. The results shown in the panels of Figure A.9 replicate those from above for the motivations: few are consistently signed in accordance with the theoretical framework, some are consistently signed the opposite of it, and few exhibit noteworthy variable importance.[36]

---

[35] We thank an anonymous reader for making this point.

[36] We should also note that the differences between these two robustness checks are small. This suggests that little of the (small, inconsistent) motivation effects on aid selectivity flow through resources. We predicted this in our speculative extension of the *motivations* approach.

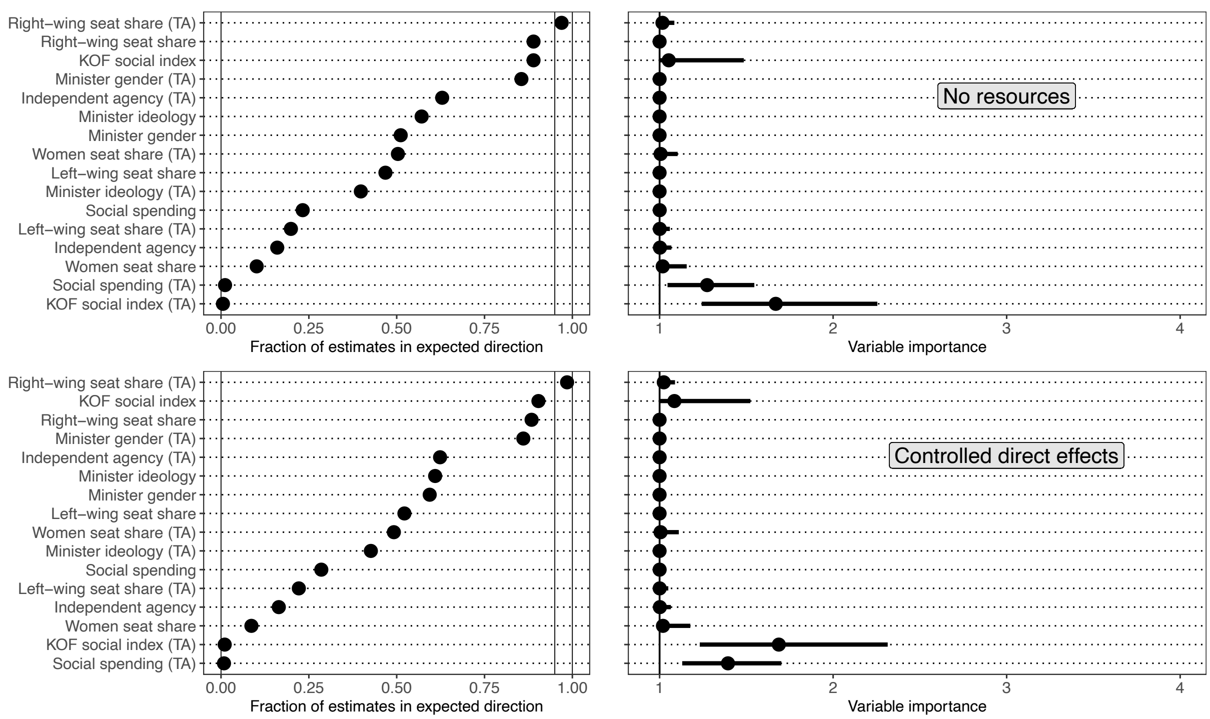

Figure A.9: Alternative estimates for motivations measures and selectivity factor. Each panel in the figure gives the fraction of non-parametric bootstrap estimates of the variables′ signs that are in the direction consistent with the theoretical framework (left side) and the corresponding variable importance (right side). The line gives the 90% confidence interval. The upper panel gives the results when dropping *donor resources* entirely and the bottom when the effect after controlling for *donor resources* (i.e. CDEs).

H Additional tables

	Commitments, Bottom 25			CPA, Bottom 25		
	Coefficient	**% Inclusion**	**% Support**	**Coefficient**	**% Inclusion**	**% Support**
Donor resources	-12.46 [-19.52; -5.13]	98.42	1.00	-12.08 [-19.52; -5.13]	99.46	1.00
Independent agency	-1.31 [-3.99; 0.02]	44.20	0.07	-0.19 [-3.99; 0.02]	16.89	0.30
Independent agency (TA)	-1.44 [-4.12; 0.01]	45.85	0.07	-0.34 [-4.12; 0.01]	21.37	0.21
KOF social index	3.32 [0.00; 8.75]	49.11	0.94	6.43 [0.00; 8.75]	82.26	1.00
KOF social index (TA)	-10.56 [-17.41; -3.85]	96.08	0.00	-15.59 [-17.41; -3.85]	99.90	0.00
Left-wing seat share	-0.09 [-0.66; 0.10]	8.70	0.36	0.40 [-0.66; 0.10]	25.96	0.92
Left-wing seat share (TA)	-0.26 [-1.62; 0.32]	19.10	0.33	-0.60 [-1.62; 0.32]	34.78	0.19
Minister gender	-0.03 [-0.22; 0.07]	5.10	0.37	-0.03 [-0.22; 0.07]	5.40	0.32
Minister gender (TA)	0.13 [-0.04; 0.80]	9.53	0.70	-0.08 [-0.04; 0.80]	9.76	0.32
Minister ideology	-0.10 [-0.66; 0.03]	9.59	0.65	-0.04 [-0.66; 0.03]	7.99	0.61
Minister ideology (TA)	0.16 [-0.05; 0.98]	12.29	0.24	-0.14 [-0.05; 0.98]	15.82	0.70
Right-wing seat share	-0.20 [-1.02; 0.02]	14.99	0.79	0.16 [-1.02; 0.02]	11.33	0.29
Right-wing seat share (TA)	-0.62 [-2.19; 0.00]	34.05	0.91	-1.67 [-2.19; 0.00]	77.82	1.00
Social spending	-0.71 [-3.25; 0.07]	20.30	0.16	0.35 [-3.25; 0.07]	12.05	0.69
Social spending (TA)	-5.10 [-9.65; -0.49]	77.27	0.01	-7.65 [-9.65; -0.49]	98.26	0.00
Women seat share	-0.79 [-3.47; 0.34]	29.14	0.20	-1.00 [-3.47; 0.34]	28.32	0.15
Women seat share (TA)	-0.50 [-4.69; 3.29]	35.82	0.38	2.70 [-4.69; 3.29]	58.73	0.91

Table A.1: Estimates for Additional Models. Each number gives the estimate for the quantity shown in the column; the range below the 95% confidence interval. Quantities summarize 1,000 BMA models based on non-parametric bootstrap draws.

	Disbursements, Bottom 25			Commitments, Percentile		
	Coefficient	**% Inclusion**	**% Support**	**Coefficient**	**% Inclusion**	**% Support**
Donor resources	-10.67	93.94	0.98	1.44	76.72	0.37
	[-19.47; -1.63]			[-19.47; -1.63]		
Independent agency	-0.86	29.39	0.15	-1.82	50.55	0.05
	[-3.65; 0.19]			[-3.65; 0.19]		
Independent agency (TA)	0.35	16.99	0.61	2.48	64.59	0.98
	[-0.66; 2.68]			[-0.66; 2.68]		
KOF social index	2.74	37.23	0.88	0.53	26.98	0.61
	[-0.16; 9.39]			[-0.16; 9.39]		
KOF social index (TA)	-6.75	76.54	0.02	-1.61	41.25	0.22
	[-14.64; -0.22]			[-14.64; -0.22]		
Left-wing seat share	-0.35	16.60	0.25	-0.18	15.58	0.35
	[-1.86; 0.08]			[-1.86; 0.08]		
Left-wing seat share (TA)	-0.10	22.80	0.48	-0.24	21.43	0.34
	[-1.84; 1.31]			[-1.84; 1.31]		
Minister gender	-0.20	13.54	0.06	0.09	8.87	0.67
	[-1.04; 0.00]			[-1.04; 0.00]		
Minister gender (TA)	-0.12	7.81	0.22	0.81	46.74	0.98
	[-0.70; 0.03]			[-0.70; 0.03]		
Minister ideology	-0.05	5.40	0.57	-0.03	6.43	0.52
	[-0.34; 0.04]			[-0.34; 0.04]		
Minister ideology (TA)	-0.63	30.82	0.92	0.20	13.66	0.20
	[-2.22; 0.00]			[-2.22; 0.00]		
Right-wing seat share	-0.64	29.35	0.89	-0.44	29.98	0.94
	[-2.33; 0.00]			[-2.33; 0.00]		
Right-wing seat share (TA)	-0.58	26.05	0.84	-0.73	40.80	0.94
	[-2.50; 0.06]			[-2.50; 0.06]		
Social spending	-0.33	17.03	0.33	0.66	21.29	0.72
	[-2.11; 0.61]			[-2.11; 0.61]		
Social spending (TA)	-3.36	63.76	0.05	-4.19	80.91	0.02
	[-7.91; 0.00]			[-7.91; 0.00]		
Women seat share	-3.68	52.62	0.04	0.55	20.91	0.75
	[-9.55; -0.02]			[-9.55; -0.02]		
Women seat share (TA)	4.57	52.34	0.88	-1.18	42.74	0.28
	[-0.27; 12.62]			[-0.27; 12.62]		

Table A.2: Estimates for Additional Models. Each number gives the estimate for the quantity shown in the column; the range below the 95% confidence interval. Quantities summarize 1,000 BMA models based on non-parametric bootstrap draws.

	Disbursements, Percentile			CPA, P'tile		
	Coefficient	**% Inclusion**	**% Support**	**Coefficient**	**% Inclusion**	**% Support**
Donor resources	-2.23	79.44	0.72	-12.35	99.17	1.00
	[-8.45; 3.87]			[-8.45; 3.87]		
Independent agency	-1.85	60.53	0.01	-1.23	38.07	0.05
	[-3.93; -0.04]			[-3.93; -0.04]		
Independent agency (TA)	3.98	92.77	1.00	3.61	86.69	1.00
	[1.50; 6.08]			[1.50; 6.08]		
KOF social index	-0.11	24.33	0.49	0.39	20.81	0.58
	[-2.00; 1.45]			[-2.00; 1.45]		
KOF social index (TA)	1.31	43.31	0.82	-5.00	71.44	0.02
	[-0.38; 4.61]			[-0.38; 4.61]		
Left-wing seat share	-0.38	24.56	0.19	1.27	61.03	1.00
	[-1.67; 0.03]			[-1.67; 0.03]		
Left-wing seat share (TA)	-0.42	32.34	0.26	-2.77	82.47	0.01
	[-2.02; 0.45]			[-2.02; 0.45]		
Minister gender	-0.02	3.40	0.28	0.11	10.49	0.76
	[-0.12; 0.02]			[-0.12; 0.02]		
Minister gender (TA)	0.16	12.96	0.76	1.40	68.54	0.98
	[-0.01; 0.83]			[-0.01; 0.83]		
Minister ideology	0.02	4.19	0.38	0.04	8.69	0.39
	[-0.04; 0.16]			[-0.04; 0.16]		
Minister ideology (TA)	-0.35	30.00	0.90	-0.20	18.09	0.72
	[-1.31; 0.00]			[-1.31; 0.00]		
Right-wing seat share	-0.58	37.19	0.94	0.56	28.36	0.15
	[-1.89; 0.00]			[-1.89; 0.00]		
Right-wing seat share (TA)	-1.22	62.82	0.99	-3.69	97.43	1.00
	[-2.58; -0.03]			[-2.58; -0.03]		
Social spending	0.80	28.44	0.84	2.21	42.17	0.93
	[-0.11; 3.22]			[-0.11; 3.22]		
Social spending (TA)	-1.64	45.48	0.13	-7.33	95.06	0.00
	[-5.08; 0.19]			[-5.08; 0.19]		
Women seat share	-1.16	39.03	0.07	0.23	18.34	0.65
	[-3.57; 0.02]			[-3.57; 0.02]		
Women seat share (TA)	0.19	30.88	0.47	2.10	50.35	0.89
	[-2.53; 3.54]			[-2.53; 3.54]		

Table A.3: Estimates for Additional Models. Each number gives the estimate for the quantity shown in the column; the range below the 95% confidence interval. Quantities summarize 1,000 BMA models based on non-parametric bootstrap draws.