

# Pre-Analysis Plan

Drivers of Distributive Preferences: Material Self-Interest vs. Expressive Benefit in Africa

*Ethan Christensen & Jennifer Hamilton*

*May 11, 2018*

## Introduction

Classic works on the motivations of voters such as Downs's 1957 classic *An Economic Theory of Democracy* and the Meltzer-Richard model hold the key assumption that citizens act rationally (1957, 36; 1981). Citizens cast votes aimed at maximizing their benefits from government, while Downs allows for utility derived from altruistic action, much research on voting behavior continues to conceptualize voting as a rational and self-interested act. These models hold, for example, that voters consider their own place in the income distribution and vote to establish a taxation level that maximizes private material self-interest (Meltzer and Richard 1981).

This idealized notion of a rational and materially self-interested voter, however, is far from what we can empirically infer about voters. Automatic emotional responses are now understood to play an important role in information processing and decision making. For instance, Charles S. Taber and Milton Lodge have collected evidence of "hot cognition," wherein political stimuli automatically and rapidly trigger an affective charge from long-term memory; these affective charges influence subsequent cognition (Lodge and Taber 2005; Morris et al. 2003). Such affectively charged cognitions fall short of the idealized rational cognition pervasive in political science models.

Recognizing the role of affect in preference formation, David Sears and coauthors developed a theory of symbolic politics (Sears et al. 1980; Sears, Hensler, and Speer 1979; Lau, Brown, and Sears 1978). Under this theory, citizens form predispositions toward political objects or symbols early in life that remain relatively stable over a lifetime. Later in life, exposure to stimulus reminiscent of that political symbol triggers an affectively charged response based on longstanding predispositions. The need for cognitive consistency outweighs present material self-interest. As a result, "Political attitudes... are formed mainly in congruence with long-standing values about society and the polity, rather than short-term instrumentalities for satisfaction of one's current private needs" (1980, 671). As a classic example, individuals most involved in anti-Vietnam war protest did not have much literal personal skin in game. They came from populations protected from the draft – young women, as well as young men who received higher education exemptions from the draft. Symbolic predispositions better explained their actions than material self-interest.

This concept of symbolic politics fits into a broader theory of expressive politics. In this model, individuals don't just have preferences about policies, they also have preferences about their preferences (Brennan and Lomasky 1993). These meta-preferences can come from symbolic sources such as those discussed by Sears and coauthors, but can also come from general group socialization. The act of voting or of expressing one's opinion is not just an opportunity to indicate one's preferences, but an opportunity to express who one is (Schuessler 1993). While these sorts of preferences can in many cases follow in the same direction as the material preferences of the Downs or Meltzer-Richard model, they can also lead individuals to act against self-interest.

While the literature on American voting behavior has sought to deal with the influence of expressive benefits, Voting behavior in sub-Saharan Africa is often focused on instrumental ethnic identity. Where ethnicity falls in the material self-interest versus expressive benefits debate is unclear. On the one hand, identity is often central to the notion of expressive politics (for instance D. L. Horowitz 1985). On the other hand, the instrumental view of ethnic voting suggests that citizens elect coethnics because it is welfare enhancing when the norm is to distribute goods along ethnic lines (Bates 1982). This line of thinking suggests that material self-interest drives ethnic voting. Several recent studies tend to support this instrumental view of ethnic

voting. For instance, recent research suggests that African citizens do consider competence or incumbent performance in their voting decisions (K. Ferree and Horowitz 2010; Bratton, Bhavnani, and Chen 2012; Carlson 2015). Ferree argues that, in South Africa, racial voting patterns arise from voters using race as a heuristic to assess parties, not from expressive voting (K. E. Ferree 2006). Moreover, evidence that region is more important than ethnicity (ICHINO and NATHAN 2013; Basedau and Stroh 2012; Ahlerup and Isaksson 2015) and that social and political institutions can affect the extent of ethnic voting (Posner 2005; Dunning and Harrison 2010) would also tend to suggest that ethnic identities are not so deep-rooted as to be a primordial, identity-based, and expressive political response. On balance, the evidence would tend to suggest that material concerns are powerful drivers of voting behavior.

Our research differs from most of this pre-existing work in several ways. First, we seek to explain policy preferences in specific domains rather than voting behavior. Second, our discussion of material self-interest versus expressive politics eschews ethnic identity, as our data is insufficient to parse whether ethnic patterns are instrumental or expressive. Instead, we focus on predominantly on political ideology, which is less explored in the African context. Some of our analysis also focuses on xenophobic tendencies. There is a substantial literature on xenophobia in South Africa, but it is largely dominated by other academic fields (see, among others, Steinberg 2012; Tella 2016; Hickel 2014; Klotz 2016; Ejoke and Ani 2017; Vromans et al. 2011).

## Description of data

In this study we plan to utilize data from the World Values Survey and Afrobarometer surveys. The World Values Survey (WVS) is the largest and longest-running non-commercial cross-national survey of beliefs and values. The most recently completed wave surveyed participants from almost 100 countries in 2011-2014. Each country-level survey has a minimum of 1,200 respondents (ages 18-85), conducted through face-to-face interviews. Depending on the country, sampling occurs through either probability or a combination of probability and stratified random sampling as noted in the documentation. From these countries, we identify the five that are in sub-Saharan Africa: Ghana, Nigeria, Rwanda, South Africa, and Zimbabwe.

Afrobarometer is a long running survey designed specifically to understand political attitudes and behaviors in sub-Saharan Africa. Respondents answer questions on a variety of topics including public services, governance, identity, and political participation. There have been six completed waves, with the most recent in 2014/2015 also covering the most countries (36). Most country surveys have either 1,200 or 2,400 respondents. Each survey uses a clustered, stratified, multi-stage, area probability sample. Occasionally a survey will oversample a politically relevant sub-population in order to ensure a large-enough subsample.

From these two sources we identify the three policy preference choices available to us: attitudes on immigration, privatization, and school fees. In order for a policy preference question to be included in our analysis, the relevant survey also had to contain appropriate items to identify the self-interest and symbolic attitudes of respondents, hence why so few policy preference items were included.

For each topic we use different samples of countries based on the data available. The relevant questions on immigration and privatization are contained in the WVS and so is limited to those five countries. We limit our analysis of education preferences to countries where there is not compelling evidence of ethnic preference in education outcomes. We choose to do so in order to focus on the differences between material self-interest and political ideology, where the nature of ethnicized preferences as either instrumental or expressive is unclear. We thus select countries for inclusion using Franck and Rainer's results on systematic bias in primary education attainment and literacy (2012). We limit our analysis to only countries in which no more than one of their four indicators attained statistical significance. This leaves us with five countries: Benin, Malawi, Mali, Senegal, and Uganda.

# Hypotheses

$H_1$ : Expressive benefit indicators will be jointly significant in each model.

$H_2$ : Material self-interest indicators will be jointly significant in each model.

## Variable construction

### Immigration

*Source*: World Values Survey Round 6

*Countries included*: Ghana, Nigeria, Rwanda, South Africa, and Zimbabwe

#### Policy Preference

(V46) When jobs are scarce, employers should give priority to people of this country over immigrants.

- *Original coding*: 1 = agree, 2 = neither, 3 = disagree, -5 = missing, -4 = not asked, -3 = not applicable, -2 = no answer, -1 = don't know
- *Recode*: 1 = agree, 2 = neither, 3 = disagree, else NA

```
wvs$V46[wvs$V46 > 0] <- NA
```

#### Expressive Benefit

1. (V107) How much you trust: People of another nationality.

- *Original coding*: 1 = trust completely, 2 = trust somewhat, 3 = do not trust very much, 4 = do not trust at all, -5 = missing, -3 = not applicable, -2 = no answer, -1 = don't know.
- *Recode*: 1 = trust completely, 2 = trust somewhat, 3 = do not trust very much, 4 = do not trust at all, else NA. *Variable type*: discrete.
- *Expected sign*: negative.

2. (V16) Important child qualities: tolerance and respect for other people.

- *Original coding*: 1 = mentioned, 2 = not mentioned, -4 = not asked, -3 = not applicable, -2 = no answer, -1 = don't know.
- *Recode*: 1 = mentioned, 0 = not mentioned, else NA. *Variable type*: binary.
- *Expected sign*: negative.

```
#Edit V107
wvs$V107[wvs$V107 < 0] <- NA

#Edit V16
wvs$V16[wvs$V16 < 0] <- NA
wvs$V16[wvs$V16 == 2] <- 0
```

#### Self-Interest

1. *Personal Connection to Immigrant*: We will create a single indicator for personal connection to immigrants using the following items. If any of these items is answered as yes (1 = immigrant), personal connection to immigrant will be coded as one, else zero:

- (V243) Mother is an immigrant.
- (V244) Father is an immigrant.

- (V245) Respondent is an immigrant.

We chose to fold these items into a single index in order to overcome potential multicollinearity. *Variable type*: binary.

*Expected sign*: negative.

2. *Employment status of primary household breadwinner* This indicator requires use of several WVS items:

- (V229) Employment status *Original coding*: 1 = full time, 2 = part time, 3 = self-employed, 4 = retired, 5 = housewife, 6 = students, 7 = unemployed, 8 = other, -4 = not asked, -3 = not applicable, -2 = no answer, -1 = don't know.
- (V235) Are you the chief wage earner in your house? *Original coding*: 1 = yes, 2 = no, -4 = not asked, -3 = not applicable, -2 = no answer, -1 = don't know.
- (V236) Is the chief wage earner employed now or not. *Original coding*: 1 = yes, 2 = no, -4 = not asked, -3 = not applicable (Respondent is the chief wage earner), -2 = no answer, -1 = don't know.

We will code 1 if either the primary wage earner of the respondent's household (either respondent or otherwise) is employed (full time, part time, or self-employed), else 0. *Variable type*: binary. *Expected sign*: positive.

3. *Supplementary analysis*: We will include the following items in supplementary analysis. However, we exclude them from our preferred model because they are subjective measures of personal well-being, and therefore subject to endogeneity vis-a-vis value expression.

(a) (V59) Satisfaction with financial situation of household.

- *Original Coding*: scale from 1 = completely dissatisfied to 10 = completely satisfied, -4 = not asked in survey, -3 = not applicable, -2 = no answer or refused to answer, -1 = don't know.
- *Recode*: scale from 1 to 10 as above, else NA. *Variable type*: discrete.
- *Expected sign*: positive.

(b) (V181) Worries: Losing my job or not finding a job.

- *Original Coding*: 1 = very much, 2 = a great deal, 3 = not much, 4 = not much at all, -4 = not asked in survey, -3 = Inapplicable (neither has, nor seeks a job), -2 = no answer, -1 = don't know.
- *Recode*: 1 = very much, 2 = a great deal, 3 = not much, 4 = not much at all, else NA. *Variable type*: discrete.
- *Expected sign*: positive.

```
#Create immigrant index variable
wvs$Immigrantindex <- ifelse(V243 == 1 | V244 == 1 | V245 == 1, 1, 0)

#create the employment index variable
wvs$Primewageemploy <- ifelse(V229 %in% c(1,2,3) | V236 ==1,1,0)
```

## Controls

1. (V240) Sex, as coded by interviewer.

- *Original coding*: 1= male, 2= female, -5 = missing/unknown, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
- *Recode*: 1 = male, 0 = female, else NA. *Variable type*: binary.

2. (V241) Year of birth

- *Original coding*: decade as indicated, -5 = missing/unknown, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: 1 = 1900-1909, 2 = 1910-1919, 3 = 1920-1929, 4 = 1930-1939, 5 = 1940-1949, 6 = 1950-1959, 7 = 1960-1969, 8 = 1970-1979, 9 = 1980-1989, 10 = 1990-1999, 11 = 2000-2010, else NA. *Variable type*: categorical.
3. (V248) What is the highest educational level that you have attained?
- *Original coding*: 1= no formal education; 2= incomplete primary school; 3= complete primary school; 4= incomplete secondary school: technical/vocational type; 5= complete secondary school: technical/vocational type; 6= incomplete secondary school: university-preparatory type; 7= complete secondary school: university-preparatory type; 8= some university-level education, without degree; 9= university-level education, with degree; -5 = refused to answer; -4 = not asked; -3 = not applicable; -2 = no answer; -1 = don't know.
  - *Recode*: 1= no formal education; 2= incomplete primary school; 3= complete primary school; 4= incomplete secondary school: technical/vocational type; 5= complete secondary school: technical/vocational type; 6= incomplete secondary school: university-preparatory type; 7= complete secondary school: university-preparatory type; 8= some university-level education, without degree; 9= university-level education, with degree; else NA. *Variable type*: categorical.
4. (V254) Ethnic group, as coded by interviewer. As indicated. *Variable type*: categorical.
5. (V24) Most people can be trusted.
- *Original Coding*: 1 = most people can be trusted, 2 = Need to be very careful, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: 1 = most people can be trusted, 0 = need to be very careful, else NA. *Variable type*: binary.
  - *Note*: This item is included in order to provide baseline trust in others for the item asking about trust in foreign nationals.

## Privatization

*Source*: World Values Survey Round 6

*Countries included*: Ghana, Nigeria, Rwanda, South Africa, and Zimbabwe

### Policy Preference

(V97) Private vs. state ownership of business

- *Original coding*: scale from 1 = Private ownership of business and industry should be increased to 10 = Government ownership of business and industry should be increased, -5 = missing, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
- *Recode*: scale from 1 = Private ownership of business and industry should be increased to 10 = Government ownership of business and industry should be increased, else NA.

### Expressive Benefit

1. (V96) Income inequality ideology.

- *Original coding*: scale from 1 = Incomes should be made more equal to 10 = We need larger income differences as incentives for individual effort, -5 = missing, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
- *Recode*: scale from 1 = Incomes should be made more equal to 10 = We need larger income differences as incentives for individual effort, else NA. *Variable type*: discrete.

- *Expected sign*: negative.
2. (V98) Government responsibility.
- *Original Coding*: scale from 1 = Government should take more responsibility to ensure that everyone is provided for to 10 = People should take more responsibility to provide for themselves, -5 = missing, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: scale from 1 = Government should take more responsibility to ensure that everyone is provided for to 10 = People should take more responsibility to provide for themselves, else NA. *Variable type*: discrete.
  - *Expected sign*: negative.

### Self-interest

1. (V230) Sector of employment.
- *Original coding*: 1 = government or public institution, 2 = private business or industry, 3 = private non-profit organization, 4 = autonomous/informal sector/other, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: 1 = government or public institution; 0 = private business or industry, private non-profit organization, autonomous/informal sector/other; else NA. *Variable type*: binary.
  - *Expected sign*: positive.
  - *Note*: this indicator is included on the basis of previous research on preferences over privatization [Battaglio\_self-interest\_2009].

### Controls

1. (V240) Sex, as coded by interviewer.
- *Original coding*: 1= male, 2= female, -5 = missing/unknown, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: 1 = male, 0 = female, else NA. *Variable type*: binary.
2. (V241) Year of birth
- *Original coding*: decades as indicated, -5 = missing/unknown, -4 = not asked in survey, -3 = not applicable, -2 = no answer, -1 = don't know.
  - *Recode*: 1 = 1900-1909, 2 = 1910-1919, 3 = 1920-1929, 4 = 1930-1939, 5 = 1940-1949, 6 = 1950-1959, 7 = 1960-1969, 8 = 1970-1979, 9 = 1980-1989, 10 = 1990-1999, 11 = 2000-2010, else NA. *Variable type*: categorical.
3. (V248) What is the highest educational level that you have attained?
- *Original coding*: 1= no formal education; 2= incomplete primary school; 3= complete primary school; 4= incomplete secondary school: technical/vocational type; 5= complete secondary school: technical/vocational type; 6= incomplete secondary school: university-preparatory type; 7= complete secondary school: university-preparatory type; 8= some university-level education, without degree; 9= university-level education, with degree; -5 = refused to answer; -4 = not asked; -3 = not applicable; -2 = no answer; -1 = don't know.
  - *Recode*: 1= no formal education; 2= incomplete primary school; 3= complete primary school; 4= incomplete secondary school: technical/vocational type; 5= complete secondary school: technical/vocational type; 6= incomplete secondary school: university-preparatory type; 7= complete secondary school: university-preparatory type; 8= some university-level education, without degree; 9= university-level education, with degree; else NA. *Variable type*: categorical.
4. (V254) Ethnic group, as coded by interviewer. As indicated. *Variable type*: categorical.

## School fees

Source: Afrobarometer Round 3

Countries included: Benin, Malawi, Mali, Senegal, Uganda

### Policy Preference

(Q10) Which of the following statements is closest to your view?. Choose Statement A or Statement B. A: It is better to have free schooling for our children, even if the quality of education is low. B: It is better to raise educational standards, even if we have to pay school fees.

- *Original Coding:* 1=Agree Very Strongly with A, 2=Agree with A, 3=Agree with B, 4=Agree Very Strongly with B, 5=Agree with Neither, 9=Don't Know, 98=Refused to Answer, -1=Missing Data
- *Recode:* 1=Agree Very Strongly with A, 2=Agree with A, 3=Agree with Neither, 4=Agree with B, 5=Agree Very Strongly with B, else NA.

### Expressive Benefit

1. (Q19) Let's talk for a moment about the kind of society we would like to have in this country. Which of the following statements is closest to your view?. Choose Statement A or Statement B. A: People should look after themselves and be responsible for their own success in life. B: The government should bear the main responsibility for the well-being of people.
  - *Original Coding:* 1=Agree Very Strongly with A, 2=Agree with A, 3=Agree with B, 4=Agree Very Strongly with B, 5=Agree with Neither, 9=Don't Know, 98=Refused to Answer, -1=Missing Data.
  - *Recode:* 1=Agree Very Strongly with A, 2=Agree with A, 3=Agree with Neither, 4=Agree with B, 5=Agree Very Strongly with B, else NA. *Variable type:* discrete.
  - *Expected sign:* negative.

### Self-interest

Afrobarometer asks two questions in separate sections regarding the affordability of school expenses and fees. We include the first in our preferred specification. However, we will run models with the second item as well for a robustness check.

1. *Public School Contact* Our self-interest items rely on having direct experience with the public schooling system. As such, we code a term for whether the respondent indicates having contact with the public school system. On each of the items listed below (Q73A-F), respondents have the option to respond "no experience with public schools in the past twelve months" (coded as 7). Our new variable is coded 0 if the respondent indicates "no experience" for ANY of the questions Q73A-F, else 1. We will interact this variable with each of the self interest variables below, effectively turning it "on" or "off" according to contact.

We do observe that the percentage of respondents choosing "no experience" sometimes vary from question to question. This is concerning, as respondents with any experience with the public school system should answer "don't know" or "never" rather than "no experience" for these questions if they have not encountered the problem listed. Because the responses of those who answer "no experience" for some questions but not others is ambiguous, we include all of these in the no contact category. *Variable type:* binary.

2. (Q73A) Have you encountered any of these problems with your local public schools during the past 12 months?. Services are too expensive / Unable to pay.
  - *Original Coding:* 0=Never, 1=Once or twice, 2=A few times, 3=Often, 7=No experience with public schools in the past twelve months, 9=Don't Know, 98=Refused to Answer, -1=Missing Data.

- *Recode:* 0=Never, 1=Once or twice, 2=A few times, 3=Often, 0=No experience with public schools in the past twelve months, else NA. *Variable type:* discrete.
  - *Expected sign:* negative.
3. *School quality index* We will construct an additive index of school quality concerns, based on the following items:
- (Q73B) Have you encountered any of these problems with your local public schools during the past 12 months?. Lack of textbooks or other supplies.
  - (Q73C) Have you encountered any of these problems with your local public schools during the past 12 months?. Poor teaching.
  - (Q73D) Have you encountered any of these problems with your local public schools during the past 12 months? Absent teachers.
  - (Q73E) Have you encountered any of these problems with your local public schools during the past 12 months?. Overcrowded classrooms.
  - (Q73F) Have you encountered any of these problems with your local public schools during the past 12 months? Poor conditions of facilities.

Each of these questions has identical response options:

- *Original coding:* 0=Never, 1=Once or twice, 2=A few times, 3=Often, 7=No experience with public schools in the past twelve months, 9=Don't Know, 98=Refused to Answer, -1=Missing Data.
- *Recode:* 0=Never, 1=Once or twice, 2=A few times, 3=Often, 0=No experience with public schools in the past twelve months, else NA.  
Response for each Q73B-F will be summed. *Variable type:* discrete.

*Expected sign:* positive.

```
af$school_contact <- ifelse(Q73A == 7 | Q73B == 7 | Q73C == 7 | Q73D == 7 | Q73E == 7 | Q73F == 7, 0, 1)
```

## Controls

- (Q90) What is the highest level of education you have completed?
  - *Original coding:* 0=No formal schooling, 1=Informal schooling (including Koranic schooling), 2=Some primary schooling, 3=Primary school completed, 4=Some secondary school/ High school, 5=Secondary school completed/High school, 6=Post-secondary qualifications, other than university e.g. a diploma or degree from a technical/polytechnic/college, 7=Some university, 8=University completed, 9=Post-graduate, 98=Refused to Answer, 99=Don't Know, -1=Missing Data.
  - *Recode:* 0=No formal schooling, 1=Informal schooling (including Koranic schooling), 2=Some primary schooling, 3=Primary school completed, 4=Some secondary school/ High school, 5=Secondary school completed/High school, 6=Post-secondary qualifications, other than university e.g. a diploma or degree from a technical/polytechnic/college, 7=Some university, 8=University completed, 9=Post-graduate, else NA. *Variable type:* categorical.
- (Q101) Respondent's gender, as assessed by interviewer.
  - *Original coding:* 1=Male, 2=Female
  - *Recode:* 1 = male, 0 = female. *Variable type:* binary.
- (Q116B) Were the following services present in the primary sampling unit/enumeration area: School?
  - *Original coding:* 0=No, 1=Yes, 9=Can't determine, -1=Missing Data.
  - *Recode Scaling:* 0=No, 1=Yes, else NA. *Variable type:* binary.



4. (Q1) How old are you?
  - *Original coding:* 18-110 as indicated, 998=Refused to Answer, 999=Don't Know, -1=Missing Data.
  - *Recode:* 18-110 as indicated, else NA. *Variable type:* discrete.
5. (Q79) What is your tribe? You know, your ethnic or cultural group.
  - *Original coding:* : see AFB for full list, 990=National identity only, 995=Other, 997=Not Asked, 998=Refused, 999=Don't know, -1=Missing Data.
  - *Recode:* As above, except NA for original codes 997, 998, 999, -1. *Variable type:* categorical.
6. (Q81) How often are [respondent's identity group] treated unfairly by the government?
  - *Original Coding:* 0=Never, 1=Sometimes, 2=Often, 3=Always, 7=Not Applicable, 9=Don't Know, 98=Refused to Answer, -1=Missing Data
  - *Recode:* 0=Never, 1=Sometimes, 2=Often, 3=Always, else NA. *Variable type:* discrete.
  - *Note:* Although we focus on countries without a history of ethnic bias in education, political preferences may nonetheless be driven by perceived ethnic bias, either in education or in other domains. As such, we include this measure to mitigate omitted variable bias vis-a-vis ethnically driven policy preference.

## Model

For each policy area, we will use an ordered logistic model with the policy preference as the dependent variable and with symbolic attitudes and self-interest indices as the independent variables. Our control variables account for respondent age, sex, education, and ethnicity in each model. We also include other control variables as listed above when we feel that they are relevant to the dependent variable.

## Immigration and Privatization Models

These models will take the form:

$$y_i = \alpha + \mathbf{X}_{i1}\beta + \mathbf{X}_{i2}\gamma + \mathbf{X}_{i3}\delta + \pi_i + \epsilon_i \quad (1)$$

where:

- $\mathbf{X}_{i1}$  represents a matrix of expressive benefit indicators and  $\beta$  represents a vector of coefficients for each expressive benefit indicator.
- $\mathbf{X}_{i2}$  represents a matrix of self-interest indicators and  $\gamma$  represents a vector of coefficients for each self-interest indicator.
- $\mathbf{X}_{i3}$  represents a matrix of control variables and  $\delta$  represents a vector of coefficients for each control variable.
- $\pi_i$  represents country fixed effects.

Note that symbolic attitudes and self-interest are thus *not* summary indices (except in cases indicated under variable construction); rather, each indicator is entered individually into the model.

*#This is the model for the immigration variable.*

```
immig_model <- MASS::polr(V46 ~ V107 + V16 + V243 + V244 + V245 + Primewageemploy + V240 + V241 + V248 +
  data = wvs,
```

```
Hess = TRUE)

#This is the model for the privatization variable.
privat_model <- MASS::polr(V97 ~ V96 + V98 + V230 + V240 + V241 + V248 + V254,
                           data = wvs,
                           Hess = TRUE)
```

## School Fees Model

This model will take the form:

$$y_i = \alpha + \mathbf{X}_{i1}\beta + \mathbf{X}_{i2} \cdot X_{i4}\gamma + \mathbf{X}_{i2}\delta + \mathbf{X}_{i4}\zeta + X_{i3}\theta + \pi_i + \epsilon_i \quad (2)$$

where:

- $\mathbf{X}_{i1}$  represents a matrix of expressive benefit indicators and  $\beta$  represents a vector of coefficients for each expressive benefit indicator.
- $\mathbf{X}_{i2}$  represents a matrix of self-interest indicators,  $X_{i4}$  represents a binary indicator for school contact, and  $\mathbf{X}_{i2} \cdot X_{i4}$  represents interactions between each self-interest indicator and school contact.  $\gamma$  represents a vector of coefficients for the interaction terms,  $\delta$  represents the vector of coefficients for self interest terms alone, and  $\zeta$  represents the coefficient of the school contact indicator alone.
- $\mathbf{X}_{i3}$  represents a matrix of control variables, and  $\theta$  represents a vector of coefficients for each control variable.
- $\pi_i$  represents country fixed effects.

Note that symbolic attitudes and self-interest are thus *not* summary indices (except in cases indicated under variable construction); rather, each indicator is entered individually into the model.

```
#here is the model for school fees.
fees_model<- MASS::polr(Q10 ~ Q19 + Q90 + Q101 + Q116B + Q1 + Q79 + Q81 + Q73A*school_contact + Q73BFIn
                        data=af,
                        Hess = TRUE))
```

## Robustness check

As an alternative to the interactive model used above, we will also conduct a robustness check in which we subset to only respondents with a 1 for school contact, and use model (1) estimate effects and test relevant hypotheses.

## Hypothesis Testing

For each model, we will conduct two likelihood-ratio tests. The first will consider whether expressive benefits indicators are jointly significant, and the second will consider whether material self-interest indicators are jointly significant.

## Expressive Benefits

The likelihood-ratio test of expressive benefits indicators has the following hypotheses:

1. For Model 1 (immigration, privatization):

- $H_0: \gamma = \mathbf{0}$
- $H_A: \beta \neq \mathbf{0}$

Hence, the restricted model will be  $y_i = \alpha + \mathbf{X}_{i1}\beta + \mathbf{X}_{i3}\delta + \pi_i + \epsilon_i$ .

2. For Model 2 (school fees):

- $H_0: \beta = \mathbf{0}$
- $H_A: \beta \neq \mathbf{0}$

Hence, the restricted model will be  $y_i = \alpha + \mathbf{X}_{i1}\beta + \mathbf{X}_{i2} \cdot X_{i4}\gamma + \mathbf{X}_{i2}\delta + \mathbf{X}_{i4}\zeta + X_{i3}\theta + \pi_i + \epsilon_i$ .

## Self-interest

The likelihood-ratio test of self-interest indicators has the following hypotheses:

1. Model 1 (immigration, privatization):

- $H_0: \beta_2 = 0$
- $H_A: \beta_2 \neq 0$

Hence, the restricted model will be  $y_i = \alpha + \mathbf{X}_{i1}\beta + X_{i3}\theta + \pi_i + \epsilon_i$ .

2. For Model 2 (school fees):

- $H_0: \gamma = \delta = \zeta = \mathbf{0}$
- $H_A: \text{at least one}\{\gamma, \delta, \zeta\} \neq 0$

Hence, the restricted model will be  $y_i = \alpha + \mathbf{X}_{i2}\gamma + \mathbf{X}_{i3}\delta + \pi_i + \epsilon_i$ .

Using the privatization model as an example, this test will be coded as:

```
p_red_exp <- MASS::polr(V97 ~ V96 + V98 + V230 + V240 + V241 + V248 + V254,
                        data = wvs,
                        Hess = TRUE)
p_red_self <- MASS::polr(V97 ~ V96 + V98 + V230 + V240 + V241 + V248 + V254,
                        data = wvs,
                        Hess = TRUE)
#Likelihood-ratio tests - pulling out the p-values.
priv_exp_p <- anova(privat_model, p_red_exp)$`Pr(>F)`[2]
priv_self_p <- anova(privat_model, p_red_self)$`Pr(>F)`[2]
```

## Standard errors

For each model, we will use cluster-robust standard errors, with primary sampling units as the cluster.

## Procedures for Missing Data

We will consider the problem of missing data separately for each model. Each question on a survey has three categories of missing data: ‘don’t know,’ ‘refused to answer,’ and ‘missing data.’ In each model, if less than 5% of observations contain any of these missing data on at least one covariate, we will use list-wise deletion and simply drop the incomplete observations. If 5% or more of observations contain missing data on at least one covariate, we will impute missing data using the **Amelia** package, using 10 imputations. While we will

use the imputed-data models when specified, we will also estimate the models using list-wise deletion models to report if imputation dramatically changes the results.

```
#####List-wise deletion#####
#Detection for the immigration and privatization models
mean(apply(wvs[,Immigration_variables], 2, anyNA)) < .05 # If TRUE, engage list-wise deletion. Else, im
mean(apply(wvs[,Privatization_variables], 2, anyNA)) < .05 # If TRUE, engage list-wise deletion. Else, im

mean(apply(afdat, 2, anyNA)) < .05 #If TRUE, engage list-wise deletion. Else, impute as below.
#####Imputation process#####
set.seed(6102018)
#impute World Values Survey (if necessary)
imputed_wvs<-Amelia::amelia(wvs,
                             m=10, #The number of imputations
                             cs="country",
                             p2s=0, #
                             ords=c("ordinals"), #specifying ordinal variables
                             noms=c("categoricals"), #specifying nominal variables
                             idvars=c("country"))

# prep storage bins for Immigration coefficients
bI.out<-NULL
# prep storage bins for Immigration standard errors
seI.out<-NULL

# prep storage bins for Privatization coefficients
bP.out<-NULL
# prep storage bins for Privatization standard errors
seP.out<-NULL

for(i in 1:imputed_data$m) { # Replicate the Immigration model on each imputed dataset
  ols.out <- MASS::polr(y ~ x1 + x2 + x3 + x4,
                       data = imputed_data$imputations[[i]],
                       Hess = TRUE)
  bI.out <- rbind(bI.out, ols.out$coef)
  seI.out <- rbind(seI.out, coef(summary(ols.out))[2])
}

for(i in 1:imputed_data$m) { # Replicate the Privatization model on each imputed dataset
  ols.out <- MASS::polr(y ~ x1 + x2 + x3 + x4,
                       data = imputed_data$imputations[[i]],
                       Hess = TRUE)
  bP.out <- rbind(bP.out, ols.out$coef)
  seP.out <- rbind(seP.out, coef(summary(ols.out))[2])
}

# Use Rubin's rules for combining results.
combined.resultsI <- mi.meld(q = bI.out, se = seI.out)
combined.resultsP <- mi.meld(q = bP.out, se = seP.out)

####impute Afrobarometer data
imputed_afdat<-Amelia::amelia(afrobarometer,
                              m=10, #The number of imputations
                              cs="country",
```

```

p2s=0, #
ords=c("ordinals"), #specifying ordinal variables
noms=c("categoricals"), #specifying nominal variables
idvars=c("country"))

# prep storage bins for School Fees coefficients
bF.out<-NULL
# prep storage bins for School Fees standard errors
seF.out<-NULL

for(i in 1:imputed_data$m) { # Replicate the School Fees model on each imputed dataset
  ols.out <- MASS::polr(y ~ x1 + x2 + x3 + x4,
    data = imputed_afdat$imputations[[i]],
    Hess = TRUE)
  bF.out <- rbind(bF.out, ols.out$coef)
  seF.out <- rbind(seF.out, coef(summary(ols.out))[2])
}

# Use Rubin's rules for combining results.
combined.resultsF <- mi.meld(q = bF.out, se = seF.out)

#Note that the likelihood-ratio tests will then also be conducted on the imputed data, requiring simila

```

## Multiple hypothesis testing

We will use the Benjamini-Hochberg Procedure to address false discovery through multiple hypothesis testing. In other words, we will:

- Order the p-values for each of  $m$  hypothesis tests such that  $p_1 \leq p_2 \leq \dots \leq p_m$ .
- Maximize  $k$  such that  $p_i \leq \frac{\alpha \cdot k}{m}$ . This value will be  $k^*$ .
- Reject all hypotheses  $H_k$  for  $k \leq k^*$

```

p <- c(p_1, p_2, p_3) #p is a vector of the p-values for all of our hypothesis tests (namely, the p-val
stats::p.adjust(p, method = "BH", n = length(p))

```

## References

- Ahlerup, Pelle, and Ann-Sofie Isaksson. 2015. "Ethno-Regional Favouritism in Sub-Saharan Africa." *Kyklos* 68 (2): 143–52. doi:10.1111/kykl.12077.
- Basedau, Matthias, and Alexander Stroh. 2012. "How Ethnic Are African Parties Really? Evidence from Four Francophone Countries." *International Political Science Review / Revue Internationale de Science Politique* 33 (1): 5–24.
- Bates, Robert. 1982. "Modernization, Ethnic Competition and the Rationality of Politics in Contemporary Africa." In *State Versus Ethnic Claims: African Policy Dilemmas*, edited by Donald Rothchild and Victor A. Olorunsola. Boulder, Colo: Westview Press.
- Bratton, Michael, Ravi Bhavnani, and Tse-Hsin Chen. 2012. "Voting Intentions in Africa: Ethnic, Economic or Partisan?" *Commonwealth & Comparative Politics* 50 (1): 27–52. doi:10.1080/14662043.2012.642121.
- Brennan, Geoffrey, and Loren Lomasky. 1993. *Democracy and Decision: The Pure Theory of Electoral Preference*. New York: Cambridge University Press.
- Carlson, Elizabeth. 2015. "Ethnic Voting and Accountability in Africa: A Choice Experiment in Uganda." *World Politics* 67 (2): 353–85. <https://muse.jhu.edu/article/580659>.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. 1st edition. Boston: Harper and Row.
- Dunning, Thad, and Lauren Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104 (1): 21–39. doi:10.1017/S0003055409990311.
- Ejoke, Ufuoma Patience, and Kelechi Johnmary Ani. 2017. "A Historical and Theoretical Analysis of Xenophobia in South Africa." *Journal of Gender, Information & Development in Africa; London* 6 (1/2): 163–85. <https://search.proquest.com/docview/1968928293/abstract/6B0F3859F5DE4BAFPQ/1>.
- Ferree, Karen E. 2006. "Explaining South Africa's Racial Census." *The Journal of Politics* 68 (4): 803–15. doi:10.1111/j.1468-2508.2006.00471.x.
- Ferree, Karen, and Jeremy Horowitz. 2010. "Ties That Bind? The Rise and Decline of Ethno-Regional Partisanship in Malawi, 1994–2009." *Democratization* 17 (3): 534–63. doi:10.1080/13510341003700394.
- FRANCK, RAPHAËL, and ILIA RAINER. 2012. "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa." *The American Political Science Review* 106 (2): 294–325.
- Hickel, Jason. 2014. "'Xenophobia' in South Africa: Order, Chaos, and the Moral Economy of Witchcraft." *Cultural Anthropology* 29 (1): 103–27. doi:10.14506/ca29.1.07.
- Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- ICHINO, NAHOMI, and NOAH L. NATHAN. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *The American Political Science Review* 107 (2): 344–61.
- Klotz, Audie. 2016. "Borders and the Roots of Xenophobia in South Africa." *South African Historical Journal* 68 (2): 180–94. doi:10.1080/02582473.2016.1153708.
- Lau, Richard R., Thad A. Brown, and David O. Sears. 1978. "Self-Interest and Civilians' Attitudes Toward the Vietnam War." *The Public Opinion Quarterly* 42 (4): 464–83.
- Lodge, Milton, and Charles S. Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26 (3): 455–82.
- Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89 (5): 914–27.
- Morris, James P., Nancy K. Squires, Charles S. Taber, and Milton Lodge. 2003. "Activation of Political Attitudes: A Psychophysiological Examination of the Hot Cognition Hypothesis." *Political Psychology* 24

(4): 727–45.

Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge ; New York: Cambridge University Press.

Schuessler, Alexander A. 1993. *A Logic Expressive Choice*. Princeton: Princeton University Press.

Sears, David O., Carl P. Hensler, and Leslie K. Speer. 1979. “Whites’ Opposition to ‘Busing’: Self-Interest or Symbolic Politics?” *The American Political Science Review* 73 (2): 369–84. doi:10.2307/1954885.

Sears, David O., Richard R. Lau, Tom R. Tyler, and Harris M. Allen. 1980. “Self-Interest Vs. Symbolic Politics in Policy Attitudes and Presidential Voting.” *The American Political Science Review* 74 (3): 670–84. doi:10.2307/1958149.

Steinberg, Jonny. 2012. “SECURITY AND DISAPPOINTMENT: Policing, Freedom and Xenophobia in South Africa.” *The British Journal of Criminology* 52 (2): 345–60.

Tella, Oluwaseun. 2016. “Understanding Xenophobia in South Africa: The Individual, the State and the International System.” *Insight on Africa* 8 (2): 142–58. doi:10.1177/0975087816655014.

Vromans, Lyn, this link will open in a new window Link to external site, Robert D. Schweitzer, this link will open in a new window Link to external site, Katharine Knoetze, and Ashraf Kagee. 2011. “The Experience of Xenophobia in South Africa.” *American Journal of Orthopsychiatry* 81 (1): 90–93. doi:http://dx.doi.org/10.1111/j.1939-0025.2010.01075.x.