
*Working Paper No. 144*

# SOCIAL DESIRABILITY BIAS AND REPORTED VOTE PREFERENCES IN AFRICAN SURVEYS

*by* **Elizabeth Carlson**

**A comparative series of national public attitude surveys on democracy, markets and civil society in Africa**

Core Partners

Center for Democratic Development (CDD-Ghana)
95 Nortei Ababio Street, North Airport Residential Area
P.O. Box LG 404, Legon-Accra, Ghana
233 21 776 142 • fax: 233 21 763 028
www.cddghana.org

The Institute for Democracy in South Africa (IDASA)
P.O. Box 56950, Pretoria Arcadia, 0007, South Africa
27 12 392 0500 • fax: 27 12 320 2414/5
www.idasa.org.za

Institute for Empirical Research in Political Economy (IREEP)
Campus ENEAM Gbegamey 02 B.P. 372
Cotonou, République du Benin
229 21 307992 • fax: 229 21 302329
www.ireep.org

with support units at

Michigan State University (MSU)
Department of Political Science
East Lansing, MI 48824
517 353 6590 • fax: 517 432 1091
www.polisci.msu.edu

University of Capetown (UCT)
Democracy in Africa Research Unit
Centre for Social Science Research
Private Bag Rondebosch 7701 South Africa
Office: 27 21 650 3827 • Dept. 27 21 650 3811
fax: 27 21 650 4657
www.cssr.uct.ac.za

www.afrobarometer.org

**AFROBAROMETER WORKING PAPERS**

*Working Paper No. 144*

# SOCIAL DESIRABILITY BIAS AND REPORTED VOTE PREFERENCES IN AFRICAN SURVEYS

*by* **Elizabeth Carlson**

**January 2014**

Elizabeth Carlson is an assistant professor of Political Science and African Studies at Pennsylvania State University.
Email: ecc13@psu.edu

**AFROBAROMETER WORKING PAPERS**

Editor
Michael Bratton

Editorial Board
E. Gyimah-Boadi
Carolyn Logan
Robert Mattes
Leonard Wantchekon

Afrobarometer publications report the results of national sample surveys on the attitudes of citizens in selected African countries towards democracy, markets, civil society, and other aspects of development. The Afrobarometer is a collaborative enterprise of the Centre for Democratic Development (CDD, Ghana), the Institute for Democracy in South Africa (IDASA), and the Institute for Empirical Research in Political Economy (IREEP) with support from Michigan State University (MSU) and the University of Cape Town, Center of Social Science Research (UCT/CSSR). Afrobarometer papers are simultaneously co-published by these partner institutions and the Globalbarometer.

Working Papers and Briefings Papers can be downloaded in Adobe Acrobat format from www.afrobarometer.org.


**Idasa**




**co-published with:** GLOBALBAROMETER

**ABSTRACT**

Much of what we know about voting behavior in Africa is based on data from public opinion surveys. However, there has been little investigation into whether reported voter preferences are reliable, or whether they are affected by bias, particularly that which may arise from the social undesirability of "tribalistic" voting. I use a voting simulation experiment in Uganda and analysis of existing surveys from a number of African countries to show that voters who are observed by others are less likely to report a preference for a coethnic candidate. I also show that coethnics of the incumbent are less likely than non-coethnics to fail to report a preference at all. Together, these findings suggest that African respondents intentionally hide their ethnic vote preferences. To measure the magnitude of the bias, I generate a counterfactual variable that estimates coethnic voters' true, but unstated, preferences. The analysis suggests that approximately 15% of survey respondents must hide an ethnic preference in order to produce significantly attenuated estimates of the effect of coethnicity on the vote.

## INTRODUCTION

A discussion among three Kenyans, captured by a Washington Post journalist in the run-up to the 2007 Kenyan election, illustrates a curious pattern in African politics. Voters consistently vote along ethnic lines, but report that ethnicity does not determine their vote.

> Along a crowded sidewalk in [Nakuru, Kenya], Esther Thuo and a friend were dis- cussing their choices in Kenya's upcoming presidential election. Thuo, a young professional, said she'd vote to keep President Mwai Kibaki in power – "Let him finish the job", she was saying – when a street vendor began heckling her.
> "You're supporting him because you're Kikuyu" – Kibaki's tribe, Peter Ambobo said.
> "I'm a Luhya," Thuo corrected him hotly, then scolded the man. "You're being tribal."
> Ambobo straightened in his chair and explained why he'd vote for fiery opposition leader Raila Odinga.
> "Not because I'm Luo" – also Odinga's tribe, he said. "I'm supporting him because I want change."
> "Okay, I'm Kikuyu," said Sammy Mbugua, a taxi driver, acknowledging the obvious explanation for why he's supporting Kibaki. "But if Kibaki goes back, he will continue his projects, like free secondary education, and this is very attractive to me."
> (McCrummen, 2007)

Two out of the three participants in the discussion indicated that they would vote for the candidate of their own ethnic group. (Thuo, the Luhya, did not have a coethnic candidate in the race.) But though straightforward ethnic preferences might seem to be the most obvious explanation for these voters' choices, all three firmly reject this explanation. Thuo is offended at the suggestion, while both Mbugua and Ambobo insist they support their coethnic candidates for decidedly non- ethnic reasons: Mbugua because he approves of Kibaki's performance and Ambobo because he does not. Though candidate performance is an entirely reasonable explanations for these voters' preferences, it may or may not be the whole story. Both men have an interest in demonstrating that they eschew blind ethnic loyalty, lest they too be scolded for being "tribal".

If ethnic preferences, or the appearance of them, are indeed socially sanctioned, this poses a problem for those who study political behavior in Africa. Following the Third Wave of democratization, a wealth of research studies have sought to identify the determinants of voter behavior throughout the continent. Partly due to the advent of the comprehensive, multi-country, multi-wave Afrobarometer survey, and partly because of a shortage of objective measures of voting behavior, most of this work has relied on survey data. Reported voting behavior or vote intentions serve as the dependent variable in studies on the effects of a range of theoretically important independent variables: actual and perceived government performance (Ferree and Horowitz, 2007; Hoffman et al., 2007; Gutierrez-Romero, 2010; Bratton, Bhavnani and Chen, 2011; Long and Hoffman, forthcoming); ideology and party (Conroy-Krutz and Lewis, 2011); clientelism (Keefer, 2010; Young, 2009; Lindberg and Weghorst, 2010); and vote buying or intimidation (Kramon, 2009; Bratton, 2008). These data have also been used to predict the winners of elections and to estimate the percent of official vote share determined by fraud (Afrobarometer, 2010; Long and Gibson, 2009). However, despite their widespread use, few studies have been done on the reliability of survey responses for measuring actual vote preferences in African countries.

Of particular concern is whether sensitivity around ethnicity in Africa produces bias in the same way that sensitivity around race does in the United States. In the US, voters tend to prefer candidates of their own race. Minority candidates are unlikely to win, for example, unless they are running in majority-minority districts (Barker, Jones and Tate, 1999; Canon, 1999; Walton and Smith, 2000). However, since voting

against someone of another race may be perceived as racist, white survey respondents who actually support a white candidate may report positive evaluations of, or even an intention to vote for, a non-white candidate (Finkel, Guterbock and Borg, 1991). Alternatively, those who oppose non-white candidates may simply avoid reporting a preference (Berinsky, 1999; Reeves, 1997). In both cases, survey results would produce overestimates of true support for non-white candidates, and underestimates of the correlation between race and vote choice.[1]

But whether ethnic voting is as sensitive for Africans as it is for Americans is an open question. The most visible proponents of "anti-tribalism" are politicians and the media, whose attitudes may be quite disconnected from the general public. African presidents routinely call for inclusivity in politics, but evidence suggests this is largely a strategy to court the votes of the ethnic opposition. In the recent Zambian presidential election, for example, President Sata campaigned on the claim that he does not believe in tribal politics, and urged voters not to vote along ethnic lines (Mulowa, 2011). Nevertheless, Sata won due largely to overwhelming support in his ethno-regional stronghold, and has since staffed the cabinet with members from the same region. Sata's call for inclusivity appears to be disingenuous. Similarly, Horowitz (2009) and Posner (2005) demonstrate that messages of inclusivity are targeted primarily to candidates' non-coethnics, which suggests that these appeals are part of an electoral strategy rather than expression of a true social norm.

Calls to avoid tribalism also dominate newspaper editorials. Recent headlines read, "Tribalism and Ethnicity Could Plunge Ghana into Anarchy" (Kunateh, 2009); "Tribalism has no Place in Namibia" (Jochem, 2007); "[Ugandan Trade Minister] Fires Salvo on Tribalism, Calls Perpetrators 'Stupid'"; and "Tribalism – A Barrier to Progressive Democratic Dialogue" (Odunze, 2013). Though presumably less politically motivated than similar statements by politicians, these opinions may be particular to the educated, wealthy, employed, and politically engaged class represented by journalists. Robinson (2009), for example, shows that education, wealth and employment are positively correlated with the likelihood that a respondent will deprioritize ethnicity in favor of other identities. This suggests that journalists may have stronger distaste for ethnic politics than the public as a whole.

Most importantly, for many African voters, ethnicity serves as an obvious target for political coordination. Voters, clustered geographically with their coethnics, have an incentive to coordinate on a coethnic candidate in expectation of receiving local development goods in exchange for their bloc vote (Ichino and Nathan, 2013; Carlson, 2013; Wantchekon, 2003). Accordingly, the community has an incentive to punish those who defect, and communities can and have resorted to social sanctions and even violence to enforce the ethnic bargain (Horowitz, 1985; Rubongoya, 2007). For most African voters, showing a preference for one's coethnic, rather than being shameful, may be socially desirable option, particularly among one's coethnics.

Nevertheless, several patterns in the data suggest that African voters consider overt ethnic preferences to be undesirable. When asked directly, voters in many countries say they do not vote ethnically. Posner (2004b) finds that a substantial number of Zambians report that ethnic voting is wrong. Lindberg and Morrison (2008) ask Ghanaians how they decide whom to vote for, and find that the vast majority report non-ethnic motivations. Bratton and Kimenyi (2008) find that a similarly low percentage of Kenyans report that they vote for candidates due to their ethnicity (even though most believe that other Kenyans do vote this way.) Nevertheless, in all three of these countries, coethnicity remains a strong determinant of

[1] This is sometimes called the "Bradley Effect" or the "Wilder Effect" after two African-American gubernatorial candidates in the 1980's earned far fewer votes than was anticipated based on polling results. It was hypothesized that support in the polls was artificially inflated by white respondents who were unwilling to report to an interviewer that they did not support a black candidate (Hopkins, 2009; Reeves, 1997).

vote choice after non-ethnic considerations such as perceived or actual economic performance are controlled for (Posner and Simon, 2002; Posner, 2004b; Bratton, Mattes and Gyimah-Boadi, 2004; Bratton and Kimenyi, 2008; Bratton, Bhavnani and Chen, 2011). This discrepancy is hard to reconcile if respondents are not systematically under-reporting the importance of ethnicity in their decision-making.

If survey respondents are unwilling to report a pure ethnic preference, analyses will end up underestimating the impact of coethnicity in predicting voter choice. For example, recent findings that presidential performance is a more important predictor of vote choice than is ethnicity (Bratton and Kimenyi, 2008; Bratton, Bhavnani and Chen, 2011; Long and Hoffman, forthcoming), may simply reflect the fact that retrospective voting is a more socially acceptable reason to prefer a candidate than is his ethnicity. Additionally, the common finding that ethnicity is not correlated to electoral outcomes in some African countries, such as Senegal or Burkina Faso (Bratton, Mattes and Gyimah-Boadi, 2004; Posner, 2004a; Basedau and Stroh, 2009; Huber, 2012) is potentially a finding about the sensitivity of reporting an ethnic vote, rather than about true voter preferences.

**IDENTIFYING BIAS**

In the United States and other developed countries, bias can be identified and measured by comparing reported and actual votes on candidates or issues at low levels of aggregation (Hopkins, 2009; Berinsky, 1999). In an African context, this is a far less viable strategy, because official tallies may or may not themselves reflect true votes or preferences. A discrepancy between reported and actual votes may be interpreted as bias on the part of respondents, but, equally plausibly, can also be interpreted as election fraud (Long and Gibson, 2009; Afrobarometer, 2010). Additionally, while American opinion polls take place daily, the Afrobarometer takes place every two years.[2] For any given country and survey round, the time between the survey and an election varies. A survey taken in close proximity to an election may reflect accurate preferences. On a survey far removed in time from an election, on the other hand, voters may struggle with the hypothetical nature of the question, or any number of events between the survey and election may lead to changes in preferences.

Methods for identifying bias in the African context, therefore, are limited to analyzing variation within responses. I use a survey experiment and variation on existing surveys to identify respondents' reported preferences under varying levels of social pressure: if voters are sensitive about reporting ethnic preferences, we should see that those exposed to social pressure are less likely to report a vote for a coethnic than are those who are able to report their preference privately or anonymously.[3] I also analyze variations in the responses of coethnics and those with no coethnic in the race. The latter group cannot possibly vote ethnically, and therefore have little reason to hide their vote for fear of appearing tribalistic. Accordingly, I compare patterns of non-response between coethnics and non-coethnics: if bias causes under-reporting of support for coethnic candidates, coethnics who prefer the incumbent should be more likely to hide this preference than non-coethnics with a similar preference for the incumbent.

Finally, to measure the potential impact of this bias on our estimates of the importance of ethnicity on vote choice, I generate a counterfactual measure of voting behavior that reflects the votes we would expect the incumbents' coethnics to report were they not constrained by social desirability concerns. Using simulated analyses using these adjusted votes, I show that, in order for our estimates of the impact

[2] An exception is a series of surveys in Uganda that took place immediately before and after the 2011 Ugandan election. These surveys very closely match actual election returns in the aggregate, but the comparison between reported votes and vote returns is only useful at the local level, where one can be reasonably sure of the ethnicity of those casting the ballots. Due to small local survey sample sizes, the difference between reported and actual vary widely at low of levels of aggregation.

[3] For evidence that privacy and/or anonymity reduce bias on sensitive issues, see Krysan (1998); Krysan and Couper (2003); Epstein, Barker and Kroutil (2001); Fisher (1993).

of ethnicity to be significantly attenuated, approximately 15% of Afrobarometer respondents whose true preference is for a coethnic incumbent must hide this preference.

**EXPERIMENTAL EVIDENCE OF BIAS**

I first test the impact of social desirability bias on voters' stated political preferences experimentally, by manipulating the level of privacy that voters have in reporting them. I use a choice experimental framework, which is specifically designed to measure the separate impact of each option's attributes on the choice that the respondent makes. I am interested in measuring the independent impact of candidates' coethnicity, relative to other relevant characteristics, on whether a respondent chooses that candidate instead of his opponent.[4] For most respondents, the experiment is conducted in such a way that respondents' preferences are reported secretly. A subset of respondents, however, undergo the same protocols, but must report their preference aloud. If the hypothesis is correct, and ethnic voting is sensitive to the point of bias, we should see that the estimated impact of a candidate's ethnicity on a voter's choice is significantly lower than it is when the voter is required to make their preference public.

Participants in the experiment were presented with pairs of hypothetical candidates and asked to report, which candidate they would prefer to see in office. The candidates were described in short vignettes and varied randomly along five dimensions: ethnicity, education, prior office held, record of goods provision in the prior office, and platform. In order to avoid stating outright whether the candidate was a coethnic, and therefore making the ethnic choice overt, I cued ethnicity by stating where in the country the candidate was from. Most ethnic groups in Uganda are strongly associated with a geographic area, and ethnicity is not necessarily associated with other cues, such as name or appearance. (See Figure 1 for sample vignettes and the appendix for the full list of candidates' possible characteristics, as well as a list of hometown proxies.)

On each dimension, approximately half of the candidates were assigned characteristics, such as coethnicity, advanced education, previous elected office, or a strong record of provision, that would be expected to increase their desirability to voters. However, because all characteristics were randomly assigned, candidates were unlikely to be assigned desirable traits on all five dimensions simultaneously. The voting simulation, by forcing respondents to choose between two non-ideal candidates, requires respondents to prioritize some characteristics over others. By analyzing the candidate characteristics and outcomes of thousands of contests,[5] we can directly measure the independent importance of a candidate's ethnicity in determining whether respondents prefer a particular candidate.

[4] The experiment presented here is part of a larger experimental study on Ugandan's vote preferences, described in detail in Carlson (2013). See this article for a full description and justification of the design, as well as balance tables and power analyses.

[5] There were 489 discrete candidate pairs read. Since each pair was read to multiple respondents, however, there were 2262 votes cast.

**Figure 1: Sample Candidate Pairing**

> Candidate One. The candidate is from Hoima. He has a Master's Degree in African Development. Before running for president, he was an LCI [village chief]. While he was LCI, he paved the community's major road. If elected president he promises to create new jobs.
>
> Candidate Two. The candidate is from Soroti. He has a university degree in Tourism. Before running for president, he was a Member of Parliament. If elected president he promises to improve government health care. (Note: This candidate is cued as having no record of provision, simply by leaving out mention of such provision.)

There are 73 ethnic groups recognized on the Ugandan census, 21 of which – those that comprise 1% or more of the population – were possible ethnicities for the candidates in this study. Since a non-weighted draw from 21 ethnicities would mean that fewer than 5% of candidates would share the respondent's ethnicity, I increased the power of the coethnicity treatment by weighting the draw. I gathered data about respondent ethnicity during recruitment and weighted the draw so that approximately 45% of candidates would be from the ethnic group(s) represented at the site.[6]
In the case that the locally dominant group was not one of the original 21 cues – such as the tiny Bwemba group, who inhabit a few villages in southwestern Uganda – I replaced the hometown with the name of the sample village. The median respondent was assigned two coethnic candidates out of six. Approximately 15% of respondents received no coethnic candidates at all.

I capture social desirability bias by estimating the importance of the pure ethnic cue on voters' choices when they are faced with varying levels of social exposure. Though most participants reported their votes using only secret paper ballots,[7] one third of participants were also assigned to report their votes aloud to a group of observers. If voting behavior is susceptible to social desirability bias, we should see that those who report their votes to others are less likely to consider ethnicity and more likely to consider other characteristics (or to give this appearance) when making their choice.[8]

All participants, both treatment and control, were assigned to vote in small groups, with the treatment assigned at the level of the group. Both treatment and control groups heard the same set of instructions about how to mark and cast their paper ballot. The treatment groups then received a short additional instruction that after casting their paper ballot they were to report their vote aloud. The instructions, with the treatment instruction indicated, are presented in the appendix.

Each group then heard descriptions of three pairs of candidates. A new set of six candidates was drawn for each group. After each pair of candidates was read, respondents silently marked a paper ballot indicating their preferred candidate, folded the ballot, and deposited it in a cardboard box at the front of the room. Once all the ballots were cast, those assigned to the treatment then also stated their vote aloud

[6] Respondents self-reported their ethnicity: in most cases, the distribution of ethnicity in the sample closely matched the ethnic distribution reported on the census.
[7] The ballots were identifiable in the sense that each ballot was coded and I was able to match individual respondents to their votes. They were secret in the sense that neither the enumerator nor the other respondents saw how the respondents voted, and no identifying information was gathered. Respondents would have little reason to think that their ballots could be traced to them: like the coded ballots used in the experiments, real Ugandan ballots have strings of numbers printed on them that are essentially meaningless to voters.
[8] See e.g. Bishop and Fisher (1995) for evidence that secret ballots produce more accurate measures of voting returns than oral exit polling on sensitive ballot issues.

to the enumerator and others in their voting group.[9] Everyone reported their preference in a given pairing before the next pair of candidates was read; in other words, voters in the treatment group heard how their peers voted in the first pairing before casting any more votes. Respondents were not given the option to abstain, in order to prevent those with socially undesirable preferences from simply failing to report any preference.

***Sample***

The sample included 754 respondents from 32 villages spread through three Ugandan regions. The sample villages were selected randomly from a Uganda Bureau of Statistics master list. Respondents of pre-randomized gender and age were recruited from every fifth household extending out in four directions from the center of the village. Each recruited respondent was randomly assigned one of six appointment times at which to arrive the next day; those arriving at the same appointment time were assembled into groups, which ranged in size from two to six, depending on attendance at that particular arrival time.[10] Every third group that was assembled was assigned to vote aloud. There were 129 voting groups, comprising 546 respondents, in the control condition, and 49 voting groups, comprising 208 respondents, assigned the treatment.

Table 1 shows that respondents assigned to the treatment and control groups were balanced in terms of age, gender, education, employment status, stated ethnic attachment, membership in the local ethnic majority, the likelihood of being assigned a coethnic enumerator,[11] and the ethnic diversity of the surrounding community.[12]

[9] There were also other people at the site, including waiting participants, additional enumerators conducting other parts of the study, and a foreign PI. These people were deliberately housed in a different space, out of earshot of the voting group, and to the extent possible, the number and identity of personnel were kept constant across groups. The results presented here are robust to site fixed effects, which should capture the effect of any unintentional variation in the "audience" at each site.

[10] Approximately 85% of recruited respondents participated and the vast majority arrived within 20 minutes of their assigned appointment time.

[11] We generally expect respondents to be more positive toward their own ethnic group, and less positive to non-coethnics, when speaking to an enumerator who shares that ethnicity, (Anderson, Silver and Abramson, 1988; Hatchett and Shuman, 1975; Davis, 1997; Campbell, 1981; Finkel, Guterbock and Borg, 1991) though there is some evidence in the American literature (Krysan and Couper, 2006; Cody, Davis and Wilson, 2010) and in this analysis that challenge this expectation. In addition, it is unclear whether a more positive attitude toward coethnics means respondents are now presenting their unbiased true preferences, or overstating their ethnic preference in response to social pressure to demonstrate ethnic solidarity. The impact of enumerator coethnicity is beyond the scope of this study and is being addressed elsewhere; I simply control for it.

[12] This is a Herfendahl index based on 2002 parish-level census data.

**Table 1: Balance of Respondent Characteristics Across Control and Treatment Groups**

| Variable | Control | Treatment | p-value of Δ |
|---|---|---|---|
| Age | 37 | 37 | 0.79 |
| Female | 53% | 57% | 0.29 |
| Years of Education | 6.5 | 6.0 | 0.13 |
| No formal employment | 59% | 62% | 0.61 |
| Prioritize national identity | 67% | 67% | 0.86 |
| Would marry non-coethnic | 72% | 75% | 0.33 |
| Locally dominant ethnic group | 80% | 81% | 0.90 |
| Coethnic enumerator | 32% | 31% | 0.46 |
| Diversity of local community | 0.41 | 0.42 | 0.47 |
| Rural | 79% | 86% | 0.04 |
| Voting group size | 3.4 | 4.4 | 0.00 |
| Heterogenous voting group | 36% | 46% | 0.01 |

However, there were three potentially important imbalances between the two groups resulting from variable attrition and its impact on the randomization process. There was higher attrition in urban areas, which reduced the sample of voters and the number of voting groups in these areas. The treatment was assigned to every third voting group; if an area had fewer voting groups, a treatment group was more likely to be dropped than a control group. For example, an urban might have five groups, one of which was assigned to treatment, (20-25%), while in rural areas, there might be six or seven groups, two of which 30-35% were assigned to treatment. In addition, since fewer urban respondents arrived at any given appointment time, control voting groups were smaller. Therefore, respondents assigned to vote aloud were significantly more likely to be rural and voted in larger groups. Those in the treatment groups were also more likely to vote in groups containing more than one ethnic group, simply because their larger size of the group increased the likelihood that a respondent from an ethnic minortiry would be represented.[13] This is a problem for causal inference if those in rural areas, those in larger groups, or those interacting with non-coethnics are less likely to select a coethnic candidate irrespective of social pressure.

In the following analyses, I will show that the impact of public voting is robust to controls for various respondent and context characteristics. Notably, two groups that are overrepresented in the treatment group – rural respondents and those voting in heterogenous groups – are actually significantly more likely to demonstrate an ethnic preference. To the extent that imbalance is affecting the estimated treatment effect of observation, it is attenuating the results.

***Experimental Results***

Columns One and Two of Table 2 show mixed-logit analysis of the first two rounds of the voting simulation, by treatment.[14] Mixed logit is optimal for an analysis of choice experiments of the design I use here (Train, 2003). It presents the independent impacts of each of a candidate's characteristics on the likelihood that the candidate will be the one chosen by the respondent. Mixed logit also controls for the correlation among the multiple choices made by each respondent.[15] To simplify the comparison, each of

[13] In order to avoid conflating rural sites with ethnically homogenous sites, heterogeneous rural areas were oversampled. This is why control and treatment respondent come from equally diverse communities, even though rural voters are overrepresented in the treatment group

[14] I exclude the third round of voting from this analysis due to strong order effects demonstrated in Table 2. The estimated treatment effect is similar, though smaller and less significant, when all three rounds are included.

[15] The results are robust to different specifications of random coefficients and expected distributions on these coefficients; to the analysis with conditional logit; and to analysis with simple-difference of means.

the candidates' five characteristics is collapsed into dummies (e.g. coethnic or non-coethnic, record of provision or no record). The results of the voting simulation among the control group – who cast only secret ballots – are shown in Model One. When votes are cast secretly, the candidate's ethnicity is four times as important as his performance in determining whether voters prefer him. Holding other characteristics at their mean, coethnicity increases a candidate's likelihood of winning by 12%, while a record of provision increases his likelihood of winning by only 3%.

**Table 2: Determinants of Candidate Selection, Given Voting Context**

| | Model One | Model Two | Model Three |
|---|---|---|---|
| | Secret Rounds 1+2 | Public Rounds 1+2 | Secret Round Three |
| | (Std. Err.) | (Std. Err.) | (Std. Err.) |
| Candidate is coethnic | 0.502*** | -0.057 | 0.012 |
| | (0.123) | (0.181) | (0.145) |
| Candidate has positive record | 0.151 | 0.615*** | 0.586* |
| | (0.104) | (0.234) | (0.302) |
| Candidate held prior office | 0.437*** | 0.279 | -0.060 |
| | (0.131) | (0.233) | (0.192) |
| Candidate has relevant degree | 0.159 | 0.399* | -0.405* |
| | (0.096) | (0.190) | (0.180) |
| Candidate promising education | 0.177 | 0.581*** | 0.144 |
| | (0.098) | (0.181) | (0.166) |
| N | 2156 | 820 | 1082 |
| Log-likelihood | -722.00 | -263.09 | -363.884 |
| $\chi^2_{(5)}$ | 2.52 | 2.88 | 0.9 |

Significance levels : † : 10% * : 5% ** : 1% *** : 0.1%

Mixed logit model. DV takes value one if respondent chose candidate and zero otherwise. Candidates are grouped by choice set (respondent-contest). Reported N represents number of candidates read; number of respondent-contests is half this number. Coethnicity and record variables take random coefficients. Estimated standard deviation of these coefficients is not shown.

Column Two of Table 2 presents the results among the treatment group. Compared to the control group, treatment voters' apparent decision-making calculus is reversed. Those assigned to report their vote aloud were disinclined to vote based on the candidate's ethnicity. The coefficient on ethnicity drops to near zero and becomes highly insignificant. Instead, voters appear to be making their choices based largely on whether the candidate has a prior record of goods provision. Among the treated sample, a record of provision increases a candidate's chance of winning by 12% when other variables are set at their means. For those in the treatment group, moreover, the candidate's platform and education suddenly become substantively and statistically significant predictors of the vote. In other words, when voting in front of others, voters apparently select their candidates based on everything except ethnicity.

To test the robustness of results to the imbalances noted above, I look only at coethnic candidates and their likelihood of winning. I control for respondent variables shown to be imbalanced above, as well as additional characteristics of the candidate and the challenger (these coefficients are not shown). Table 3 confirms that coethnic candidates' vote shares decrease significantly when voters are asked to share their preferences aloud controlling for imbalances between the groups. The coefficient on the public voting treatment in Table 3 translates to a 9% reduction in a coethnic's likelihood of victory, which is significant at the 0.05 level.

Table 3 demonstrates that the disproportionately rural composition of the treatment group cannot explain why the treatment group is less likely to vote ethnically, as rural voters are more likely to prefer a

coethnic. Neither is the treatment effect being driven by the more diverse groups in the treatment condition. Respondents asked to vote in a heterogenous voting group are more likely to support a coethnic than those voting only among their own coethnics. If ethnic groups form tighter social networks (Habyarimana et al., 2007), it is possible that voting publicly among coethnics constitutes more consequential social exposure and therefore a stronger treatment. A diverse group may also prime ethnicity and ethnic competition in a way that homogenous groups do not. Whatever its genesis, the results not only reduce our concern about the imbalance between control and treatment groups, but also suggests that the impetus to avoid appearing tribalistic is stronger than the incentive to coordinate on a coethnic candidate.

**Table 3: Determinants of Coethnic Candidate Victory**

| Variable | Coefficient (Std. Err.) |
|---|---|
| Public voting treatment | -0.367* (0.182) |
| Voting group size | -0.179** (0.067) |
| Heterogenous voting group | 0.429* (0.191) |
| Rural | 0.364† (0.201) |
| Intercept | 0.701* (0.343) |
| N | 701 |
| Log-likelihood | -461.317 |
| $\chi^2_{(6)}$ | 36.864 |

Significance levels : † : 10% * : 5% ** : 1%

Standard logit model. Sample is limited to coethnic candidates from first two rounds of voting. DV takes one if candidate won his contest. Includes controls for other characteristics of both candidate and challenger. Standard errors clustered by respondent.

Table 4 demonstrates that the effect of the treatment is, indeed, a result of voters' desire to avoid showing an overt ethnic preference, that is, a preference for a coethnic candidate who does not have a good record with which to justify his selection. When the vote is not observed, a coethnic without a record of provision wins 47% of the time. When the vote is public, a coethnic wins only 37% of the time – a 10 percentage-point reduction. There is a much smaller decline in the likelihood of victory for a coethnic who does have a record – the treatment reduces the likelihood such a candidate will win by only 4%. The votes lost by the coethnics without records are picked up by non-coethnics who have records: their chance of victory increases from 46% to 57% from the control to treatment group. This confirms that voters are switching away from an apparent pure ethnic preference and toward a more acceptable preference for a well-performing candidate. If respondents on surveys respond the same way, shifting their support from non-performing coethnics to well-performing coethnics, this would not only attenuate the apparent impact of coethnicity on the vote, but also exaggerate the apparent effect of performance.

**Table 4: Percent of Candidates Winning, by Characteristics and Context**

| | Non-Coethnic | | Coethnic | |
|---|---|---|---|---|
| | Secret | Public | Secret | Public |
| No Record | 0.45 | 0.43 | 0.47 | 0.37 |
| Good Record | 0.46 | 0.57 | 0.66 | 0.62 |

The sensitivity of voters to the context of the simulation indicates that respondents are quite sophisticated in their attempts to avoid appearing tribalistic. Respondents were not informed about the importance of ethnicity in the study. During recruitment, potential respondents were told that they would be taking part in a study about what Ugandan voters desire in a presidential candidate. They were not asked directly whether they prefer coethnic candidates to non-coethnics, nor to explain their preferences for certain candidates. Ethnicity was signaled indirectly, through the hometown proxy, and every candidate had three characteristics other than record with which a respondent could justify voting for him if pressed to do so. Nevertheless, respondents were able to determine the importance of ethnicity in their selection and realize that supporting a coethnic, especially over a well-performing non-coethnic, would indicate socially unacceptable preferences. This finding indicates that there is a possibility respondents feel compelled to show non-ethnic preferences on public opinion surveys, even when researchers takes pains to hide the relevance of ethnicity in their questions.

**OBSERVATION IN THE SURVEY SETTING**

The experiment demonstrates that Ugandans exposed to social pressure are substantially less likely to report support for a coethnic presidential candidate. As with all experiments, however, this one raises concerns about external validity: first, whether the experimental findings hold in a non-experimental setting, and second, whether the results can be generalized outside Uganda. In order to maximize control in the experimental setting, the candidates in the experiment were hypothetical and their descriptions sparse. In answering questions about real candidates, voters may have other information available on which to make their choice. More importantly, in a typical survey setting, there may never be a need for a respondent to demonstrate a pure ethnic preference unconnected to performance because respondents can provide their own evaluations of the incumbent's performance. Coethnics of Uganda's President Museveni overwhelmingly report that they approve of his performance, and may feel that this sufficiently justifies their support for him. To the extent that the experimental conditions carry over to the survey setting, it is also possible that the findings hold only in Uganda, where a history of major state-sponsored ethnic violence may have made ethnicity particularly sensitive.[16] In this section, I test for observation effects in survey data from Uganda and 13 additional Afrobarometer countries.

Using an unintended variation on Round Four of the Afrobarometer, in which some interviews were observed by others in the community, I provide evidence that observation in a survey context can also reduce the likelihood that voters will report a preference for a coethnic candidate. Much of African life – including interviews with researchers – takes place out of doors. Accordingly, survey enumerators are sometimes unable to keep away curious onlookers who wish to hear what their family member or neighbor is being asked. About one third Afrobarometer interviews are at least partially observed by others in the household or community, and this is noted by the enumerator. The treatment is not a strong one. Interviews are observed only by family members. Also, we only know that the respondent was observed at some point during the survey, not necessarily that they were observed while answering questions about voting behavior. This should make it more difficult to find an impact from observation on voting behavior.[17]

[16] Both Idi Amin and Milton Obote killed thousands of their rivals' coethnics.

[17] The countries used in the pooled analysis are Benin, Botswana, Burkina Faso, Ghana, Kenya, Madagascar,

Because observation by others is not randomly assigned on the Afrobarometer, however, I control for a number of characteristics correlated with both the likelihood of observation and reported vote choice: urban residence, poverty as measured by insufficient access to food, gender, age, education, membership in the country's largest ethnic group, membership in the president's ethnic group, interaction with a coethnic enumerator, and interaction with an enumerator of the president's ethnicity.

The model takes as its dependent variable a dummy indicating whether the respondent reported he would vote for a presidential candidate of his own ethnicity. This was not asked directly; rather, the question asked which party the respondent would vote for if an election "were held tomorrow." I coded as coethnic or non-coethnic according to the ethnicity of the presidential candidate from that party in the election closest in time to the survey.[18] In Model One, the sample includes only those who stated a preference. The variable takes a one if the respondent indicated that he would vote for the coethnic candidate, and a zero if he said he would vote for a candidate of another ethnicity. In Model Two, those who did not report a preference are included and coded as non-ethnic voters. To avoid inflating the pool of non-ethnic voters, those voters who do not have a coethnic option among the major parties in the race are excluded. Only coethnics of candidates earning at least 10% of the vote are included. I allow the impact of observation to have heterogenous affects across countries by including country random effects.

The model confirms that, in the survey setting, in a sample of multiple African countries, observation reduces the likelihood that a respondent who has a coethnic candidate in the race will report a preference for that candidate. Holding other variables at their means, observation reduces the likelihood of reporting an ethnic preference by approximately 10%. The results are also consistent with additional findings from the experiment presented in Table 3. Respondents in rural areas are more likely to report support for a coethnic, while those who are interviewed by a coethnic enumerator are less likely to do so. This should increase our confidence that the results of the Ugandan experiment have implications for the reliability of African survey responses in general.

***"Non-Ethnic" Countries***

In most countries in the above pooled sample, including Uganda, the effect of observation, when analyzed in a single-country sample, is not statistically significant due perhaps in part to a loss of power in the disaggregated sample. In three countries, however, the impact of observation on reported vote is large in magnitude and highly significant. In Burkina Faso, Senegal and Madagascar, respondents who are observed are ten to fifteen percent less likely to report support for a coethnic. Since the analyses in these countries are also underpowered, these large coefficients are more likely to be erroneous overestimates than they would be in a larger sample. Nevertheless, it is striking that all three countries in which we get these significant estimates are countries in which ethnicity has previously been found to be a poor predictor of vote (Bratton, Mattes and Gyimah-Boadi, 2004; Basedau and Stroh, 2009; Huber, 2012; Dunning and Harrison, 2010) and are cited by area experts as examples of states in which ethnic cleavages are minimal or even non-existent (see e.g. Fearon and Laitin, 2005; Cheeseman and Ford, 2007; Posner, 2004a). These results raise an additional concern that these earlier findings, many of which are based on survey data, are picking up, not the irrelevance of ethnicity, but its extreme sensitivity for voters, who are unwilling to report a genuinely held preference for a coethnic candidate.

---

Malawi, Mali, Mozambique, Namibia, Senegal, Uganda and Zambia. Some countries with surveys in Round Four were left out because either all candidates were of the same ethnicity, the sample of those with coethnic candidates was too small to analyze, or the salient ethnic cleavage could not be clearly identified.

[18] In Ghana, I coded the parties according to their traditional ethnic affiliation, rather than the affiliation of the current candidate(s), because the former is far more predictive of votes.

**Table 5: Likelihood of Reporting Intention to Vote for Coethnic Candidate**

| | Model One | Model Two |
|---|---|---|
| Others watching | -0.140* | -0.088† |
| | (0.066) | (0.049) |
| Rural | 0.432** | 0.256** |
| | (0.067) | (0.050) |
| Education (yrs.) | -0.020* | -0.008 |
| | (0.009) | (0.007) |
| Age | 0.000 | -0.001** |
| | (0.000) | (0.000) |
| Food shortage | 0.009 | 0.008 |
| | (0.027) | (0.020) |
| Member of dominant group | 0.740** | 0.704** |
| | (0.088) | (0.070) |
| Member of incumbent's ethnicity | 0.922** | 0.177** |
| | (0.084) | (0.067) |
| Coethnic enumerator | -0.442** | -0.258** |
| | (0.073) | (0.056) |
| Enumerator of incumbent's ethnicity | 0.251** | 0.057 |
| | (0.094) | (0.072) |
| Believes gov't sponsoring survey | 0.296** | 0.377** |
| | (0.063) | (0.048) |
| Intercept | 0.204 | -0.423† |
| | (0.380) | (0.228) |
| N | 6108 | 8959 |
| Log-likelihood | -3300.145 | -5649.431 |
| $\chi^2_{(10)}$ | 268.671 | 254.785 |

Significance levels : † : 10% * : 5% ** : 1%

In both models, the dependent variable takes a one if the respondent indicated support for their coethnic candidate. In Model One, the dependent variable takes a zero if the respondent indicated support for a non-coethnic candidate. In Model Two, the dependent variable takes a zero if the respondent reported support for a non-coethnic or did not report a prference at all. Model includes country random effects.

## SOCIAL DESIRABILITY IN AN ANONYMOUS SURVEY CONTEXT

In most surveys, of course, respondents are not observed, except by an enumerator. The enumerator has usually assured the respondent that his or her responses will remain confidential, and unlike the observers in the experiment, does not contribute his or her own views. Without the additional impact of observation, is there reason to believe that social desirability bias affects data gathered in a straightforward survey context? I do not have the ability, given data from either the experiment or the Afrobarometer survey, to determine the difference in response between those who report their vote secretly and those who report only to an enumerator. In this section, however, I provide evidence from the experiment that raises concerns for the reliability of normal survey responses, and test for evidence of response bias in existing surveys.

First, the experimental results suggest that it is the risk of being heard to voice socially undesirable choices, rather than hearing the votes of others that is driving voters to change their votes. The votes that respondents marked on the secret ballot before they heard anyone else's vote, and the vote that they reported aloud, are the same in almost all cases (over 97%). This indicates that voters anticipated having to report their vote aloud and changed their vote to the socially appropriate choice in advance. Voters did

not wait to see how others voted before making their selection. This suggests that reporting one's vote aloud is problematic, whether or not the listener(s) also report(s) a preference.

Additionally, there is evidence that ethnicity is sensitive even when participants' responses are not seen by either other voters or an enumerator as long as respondents know that they are being asked to report on their ethnic preferences. As noted, respondents heard three pairs of candidates and voted three times. The protocols for the third round of voting were identical in every sense to those of the first two rounds. By the third round of voting, however, the median respondent had encountered two coethnic candidates – far more than random chance would predict – making the role of ethnicity in the choice overt. Accordingly, as Table 2 shows, the decisions made by control and treatment groups in Round 3 are very similar. Respondents in the control group act as if they are voting publicly. Accordingly, coethnicity no longer predicts preference, even though these same voters had reported ethnic preferences in earlier rounds.[19] This suggests that, as long as voters understand the ethnic context of the question, anonymity and/or confidentially may not be enough to preserve the integrity of their answers. Notably, on the most recent rounds of the Afrobarometer survey, questions on party preference – which are strongly correlated with later questions about reported vote – follow immediately after a series of questions on ethnic identity, salience, and trust. Rather than playing down the role of ethnicity in party politics, this ordering would presumably prime it. The ordering may alert respondents that their political preferences are interpreted through an ethnic lens.

***Non-Response***

Following Berinsky (1999), we can leverage differences in non-response[20] between coethnics and non-coethnics as measure of social desirability bias. In the absence of social desirability bias, we should expect coethnics to be more likely to report a preference. The incumbent has control over government resources and is likely to distribute them to his supporters. In this scenario, respondents have an incentive to signal support for the incumbent. In addition, Weghorst (2011) has found that the relatively illiberal democratic environment in many African countries increases reported support for the incumbent because voters fear retribution for supporting the opposition. It is therefore easy to understand why someone might prefer non-response to reporting support for the opposition, but less obvious why anyone would hide genuinely-felt support for the incumbent. Since coethnics are more likely to genuinely prefer the incumbent, they should also feel less conflicted over whether they should respond, and therefore have higher response rates. However, as Table 6 shows, coethnics in the full Afrobarometer sample are actually significantly less likely to report a vote at all. Controlling for other determinants of non-response, including age, gender, education, political information, and various aspects of the survey context, and utilizing country random effects, coethnics of the incumbent are 11% less likely to report an electoral preference.

[19] The results of the first two rounds are indistinguishable from one another.

[20] This may take the form of outright refusal, indicating no preference, or giving a "don't know" response.

**Table 6: Likelihood of Reporting Any Vote Preference**

| Variable | Coefficient (Std. Err.) |
|---|---|
| Incumbent's coethnic | -0.086* (0.040) |
| Age | -0.001** (0.000) |
| Female | -0.207** (0.029) |
| Education (yrs.) | -0.005 (0.003) |
| Urban | -0.149** (0.031) |
| Believes free to express opinion | 0.020† (0.011) |
| Frequency listen to news | 0.061** (0.011) |
| Percent of questions refused | -0.223** (0.019) |
| Believes gov't sponsoring survey | 0.361** (0.029) |
| Coethnic enumerator | -0.096** (0.031) |
| Intercept | 0.786** (0.163) |
| N | 25233 |
| Log-likelihood | -14790.63 |
| $\chi^2_{(10)}$ | 500.666 |

Significance levels : † : 10% * : 5% ** : 1%

Dependent variable is a dummy indicating whether the respondent reported that they had a preference for a candidate in the upcoming election. Refusal and "don't know" are coded as zeros. Model includes country random effects.

The distribution of non-response across the coethnic sample is not uniform. Coethnics who choose not to respond to the vote preference question are significantly less likely to have reported approval of the president's performance: the difference is 14% and significant at the 0.001 level. One interpretation of this finding is that disapproval corresponds with a genuine lack of support for the president, but since coethnics face particular pressure to support the incumbent, they are less free to admit to this lack of support than are non-coethnics.[21] Another interpretion, however, is that coethnic respondents who have already reported that they do not approve of the president's performance early in the survey, realize that they cannot later admit to prefering the incumbent without also admitting a pure ethnic preference. This interpretation is consistent with the results of the experiment, which show that it is those whose coethnic candidate is not a good performer who feel most strongly the need to hide their ethnic preference. Ambiguity in the interpretation means that we cannot state for certain that coethnics who fail to respond

[21] One challenge to this interpretation is that if coethnics are expected to support the president, enough that they would hide a true preferences for the opposition, it is unclear why this would not also extend to providing positive performance evaluations.

are doing so to hide a true preference for the incumbent. It also means we cannot be confident that they are not. There is little in the data thus far to convince us that existing opinion surveys are free of systematic under-reporting of ethnic preferences.

**HOW BAD MIGHT THE BIAS BE?**

Ultimately, what we want to know is how badly social desirabililty concerns may be biasing our results, specifically how they bias our estimates of the effect of coethnicity on vote choice. In this section, I run a vote choice model on the original survey results, and also run the model using two sets of counterfactual survey results reflecting respondents' predicted vote preferences. I show that estimates of the impact of coethnicity on vote choice will begin to be significantly attenuated if 16% of coethnic voters either erroneously claim to support the incumbent or fail to report a preference at all.

To generate a counterfactual vote, I predict every respondent's true latent support for the incumbent – $\widehat{p_{inc}}$ – using non-ethnic correlates of support for the incumbent, such as education and rural residence. I also generate a model of every respondent's latent likelihood of answering the question, based on variables such as gender, political interest, and general cooperation with the survey. For each coethnic in the sample, I generate a "predicted incumbent support" dummy, which takes a value of one with probability $\widehat{p_{inc}}$. The resulting variable, the counterfactual vote, represents what a respondent's expected response would be were he or she not constrained by social desirability concerns.[22] I also generate a variable "predicted responder", which takes a one with probability $\widehat{p_{resp}}$. This represents the respondent's true propensity for reporting a preference for any candidate absent social desirability bias.

I then replace some portion of coethnics' reported votes with their predicted vote. Since we know that some voters are misreporting their preferences, but not precisely how many, I determine the percentage that would have to be misreporting their preference in order to produce biased estimates. I begin by assuming a misreporting rate of zero – meaning that all votes in the original data represent genuine preferences – and increase this rate incrementally up to 30%. With each increment, I randomly replace that proportion of reported coethnic votes with their predicted vote. In the first set of simulated data, I only replace the votes of coethnics who reported a preference in the original data, leaving missing data as missing. In the second dataset, I also replace the votes of those predicted to be "true responders", but who did not actually respond in the original data. The data are simulated, and coefficients estimated, with 50 repetitions at each increment.

The model used to generate a coefficient for the impact of coethnicity on the vote, controlling for demographic characteristics, attitudinal and contextual factors is shown in Equation 1. This same model was used with all three sets of data – the original data and the two simulated datasets. Figure 2 shows the simulated coefficient on coethnicity as the misreporting rate increases from zero to 30% in the two simulated datasets, with the lower bounds of each 95% confidence intervals. The solid horizontal line reflects the original coefficient when all responses are treated as genuine (i.e. in the original data), and the dotted horizontal line shows the upper 95% bound on that estimate. The value of interest is the rate of mispreports above which the simulated coefficient using the adjusted vote data becomes significantly different from the original. Using the simulated data in which only responders' votes are replaced, the coefficient on coethnicity becomes significantly different from the original when approximately 24% of coethnics misreport their preference. If we replace the missing votes of all likely responders, approximately 16% of coethnics who have a true preference for the incumbent must either report a vote for the opposition or fail to report a preference in order to cause a significant attenuation of the coethnicity coefficient.

[22] Neither the original vote model nor the adjusted vote model have large $r^2$, but the large amount of error is accounted for in the simulation.

**Figure 2: Simulated Coefficient vs. Original Coefficient on Coethnicity**

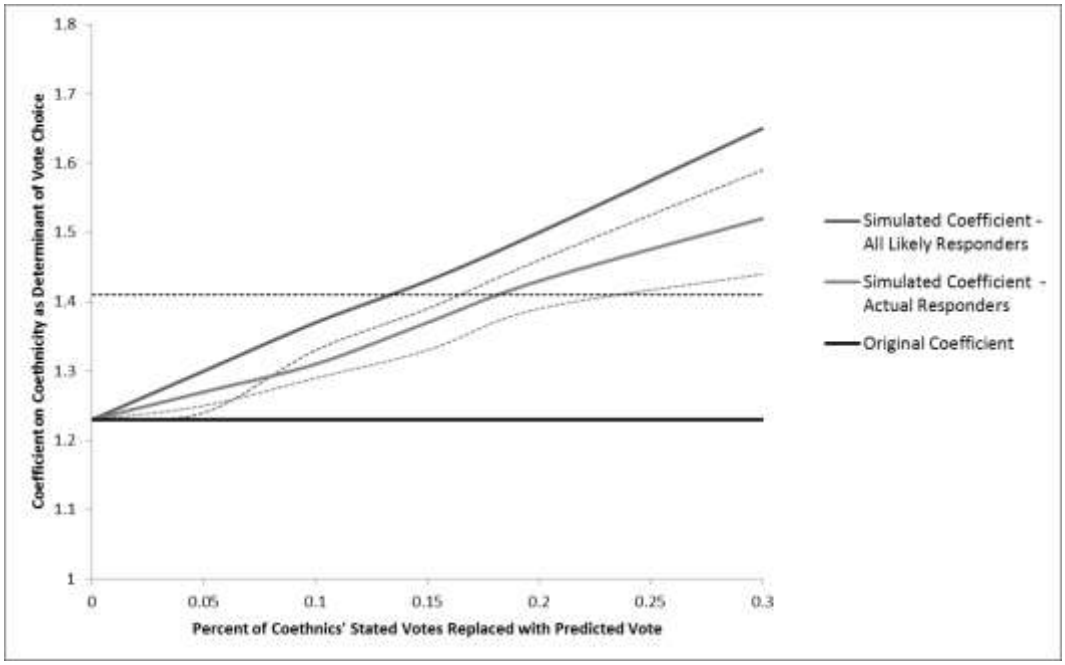


Figure represents coefficients on coethnicity, relative to support for the incumbent, from original Afrobarometer data and two simulated datasets representing coethnic voters' predicted true preferences. The solid lines represent the point estimates and the dotted lines represent the upper or lower bounds of the 95% confidence intervals of these estimates.

$$Vote_{inc} = a + b_1 Dem + b_2 Approval + b_3 Economicevaluations + b_4 Gov't Spons + \mathbf{b_5 Coethnic} + e \qquad (1)$$

Correcting social desirability bias is not enough to change our predictions of who will win or lose an African election. Even if we replace 30% of coethnic votes, overall support for the incumbent increases by only 3%. Replacing the votes of likely coethnic responders with their predicted vote, however, does change the expected marginal impact of coethnicity on determining the vote. Setting all other variables in the model at their means, a coethnic is 29% more likely to support the incumbent than is a non-coethnic when we use the original data. If we replace the votes of 16% of coethnic voters, not including missing responses, the impact of coethnicity increases to 33%. If we replace 24% of likely responders, coethnics are 35% more likely to support the incumbent.

## CONCLUSIONS AND SOLUTIONS

The results of the Ugandan experiment presented here raise serious concerns about the reliability of self-reported support for coethnic presidential candidates in Sub-Saharan Africa. A search for bias in existing survey results provides some evidence that this bias exists outside the experimental context, as well as outside Uganda. The results of this analysis suggest that voting behavior measued by Afrobarometer surveys, particularly responses that indicate ethnic preferences, are subject to under-reporting due to social desirabilty bias.

One solution is to avoid the use of surveys for studying vote choice altogether, and instead to rely on actual voting returns. However, as noted, this places an untenable constraint on researchers. In many African countries, electoral returns are only available at high levels of aggregation, which poses a serious ecological inference problem. In addition, electoral returns may reflect fraud or malfeasance and therefore may be less reflective of true preferences than are survey data. These concerns suggest that the only real solution is to retain the use of surveys, but alter standard survey protocols to allow responses on ethnic

questions to be as private as possible, without interaction with either community members or an enumerator.

These changes have potentially high costs to researchers. In the American context, subjects can have privacy by self-administering questionnaires. However, this usually involves respondents who can read and write, and a substantial minority of African respondents are illiterate. Possible alternatives include having respondents respond to enumerator questions with paper ballots like those used in the experiment presented here, or entirely self-administering parts or all of the the survey using yes/no responses to pre-recorded taped questions heard privately through earphones (e.g. Chauchard, 2010). Explaining these non-standard protocols to respondents adds time to what are already intensive surveys. However, these aformentioned protocols may be the only way to ensure reliable results on what is one of the most critical questions on the Afrobarometer survey. Before instituting these changes it may be worthwhile to gather data on voting behavior using smaller purpose-designed surveys that attempt to eliminate survey bias and then compare the results of these studies directly to existing Afrobarometer data to ensure bias is actually present in the particular country being analyzed.

For the large amount of research that has already been done using survey-based voting indicators, the evidence presented here does not call for scrapping earlier results, nor does it demonstrate that earlier conclusions are wrong. However, it does suggest that conclusions drawn about voting behavior should be taken as tentative until either the data can be shown to be unbiased or the conclusions are confirmed using alternate methods.

**REFERENCES**


Afrobarometer. 2010. "Summary of Results, Uganda 2010." Results of Round 4.5 Pre-Election Survey.

Anderson, Barbara, Brian Silver and Paul Abramson. 1988. "The Effects of the Race of the Interviewer on Race-related Attitudes of Black Respondents in a SRC/CPS National Election Studies." *Public Opinion Quarterly* 52:289–324.

Barker, Lucius, Mack Jones and Katherine Tate. 1999. *African-Americans and the American Political System.* New Jersey: Prentice Hall.

Basedau, Matthias and Alexander Stroh. 2009. "Ethnicity and Party Systems in Francophone Sub-Saharan Africa." GIGA Working Papers 100.

Berinsky, Adam. 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* 43(4):1209–1230.

Bishop, George and Bonnie Fisher. 1995. ""Secret Ballots" and Self-Reports in an Exit-Poll Experiment." *Public Opinion Quarterly* 59(4):568–588.

Bratton, Michael. 2008. "Vote Buying and Violence in Nigerian Election Campaigns." Afrobarometer Working Paper 99.

Bratton, Michael and Mwangi S. Kimenyi. 2008. "Voting in Kenya: Putting Ethnicity in Perspective." Afrobarometer Working Paper 95.

Bratton, Michael, Ravi Bhavnani and Tse-Hsin Chen. 2011. "Voting Intentions in Africa: Ethnic, Economic, or Partisan." Afrobarometer Working Paper 127.

Bratton, Michael, Robert Mattes and E. Gyimah-Boadi. 2004. *Public Opinion, Democracy and Market Reform in Africa*. New York: Cambridge University Press.

Campbell, Bruce. 1981. "Race-of-interviewer Effects Among Southern Adolescents." *Public Opinion Quarterly* 45:231–244.

Canon, David. 1999. *Race, Redistricting, and Representation: The Unintended Consequences of Black Majority Districts*. Chicago: University of Chicago Press.

Carlson, Elizabeth. 2013. "Explaining African Voters' Ethnic Preferences." Working paper.

Chauchard, Simon. 2010. "Asking Questions in an Indian Village: The Impact of an Original Audio Self-Administered Methodology on Misreporting." Working Paper.

Cheeseman, Nicholas and Robert Ford. 2007. "Ethnicity as Political Cleavage." Afrobarometer Working Paper 83.

Cody, Janay, Darren Davis and David Wilson. 2010. "Race of Interviewer Effects and Interview Clustering." Working Paper.

Conroy-Krutz, Jeffrey and Dominique Lewis. 2011. "Mapping Ideologies in African Landscapes." Afrobarometer Working Paper 129.

Davis, Darren. 1997. "The Direction of Race of Interviewer Effects among African-Americans: Donning the Black Mask." *American Journal of Political Science* 41(1):309–322.

Dunning, Thad and Jennifer Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104(1):21–39.

Epstein, Joan Faith, Peggy Ripley Barker and Larry Kroutil. 2001. "Mode Effects in Self-Reported Mental Health Data." *Public Opinion Quarterly* 65:529–549.

Fearon, James and David Laitin. 2005. "Burkina Faso." Narrative Appendix.

Ferree, Karen and Jeremy Horowitz. 2007. "Identity Voting and the Regional Consensus in Malawi." Afrobarometer Working Paper 72.

Finkel, Steven, Thomas Guterbock and Marian Borg. 1991. "Race-of-Interviewer Effects in a Pre- election Poll Virginia 1989." *Public Opinion Quarterly* 55(3):313–330.

Fisher, Robert. 1993. "Social Desirability Bias and the Validity of Indirect Questioning." *Journal of Consumer Research* 20:105–120.

Gutierrez-Romero, Roxana. 2010. "The Role of Ethnic Identity and Economic Issues in the 2007 Kenyan Elections." CSAE Working Paper Series 2010-06.

Habyarimana, James, Macartan Humphreys, Daniel Posner and Jeremy Weinstein. 2007. "Why Does Ethnic Diveristy Undermine Public Goods Provision?" *American Political Science Review* 101(4):709–725.

Hatchett, Shirley and Howard Shuman. 1975. "White Respondents and Race of Interviewer Effects." *Public Opinion Quarterly* 39(4):523–527.

Hoffman, Barak, Clark Gibson, Karen Ferree and James Long. 2007. "Explaining the African Vote." Presented to American Political Science Association.

Hopkins, Daniel J. 2009. "No More Wilder Effect, Never a Whitman Effect: When and Why Polls Mislead about Black and Female Candidates." *Journal of Politics* 71(3):769–781.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Horowitz, Jeremy. 2009. "Ethnic Groups and Campaign Targeting in Kenya's 2007 Election." Conference Paper.

Huber, John D. 2012. "Measuring Ethnic Voting: Does Proportional Representation Politicize Ethnicity?" *American Journal of Political Science* 56(4):986–1001.

Ichino, Nahomi and Noah Nathan. 2013. "Crossing the Line: Local Ethnic Geography and Voting in Ghana." *American Political Science Review* 107(2):344–361.

Jochem, Jora. 2007. "Tribalism has No Place in Namibia." *Republikein*. May 18.

Keefer, Philip. 2010. "The Ethnicity Distraction: Political Credibility and Partisan Preferences in Africa." Policy Research Working Paper 5236.

Kramon, Eric. 2009. "Vote-Buying and Political Behavior: Estimating and Explaining Vote Buying's Effect on Turnout in Kenya.'' Afrobarometer Working Paper 114.

Krysan, Maria. 1998. "Privacy and the Expression of White Racial Attitudes." *Public Opinion Quarterly* 62:506–544.

Krysan, Maria and Mick Couper. 2003. "Race in the Live and Virtual Interviews: Racial Deference, Social Desirability and Activation Effects in Attitude Surveys.'' *Social Psychology Quarterly* 66(4):364–383.

Krysan, Maria and Mick Couper. 2006. "Race of Interviewer Effects: What Happens on the Web." *International Journal of Internet Science* 1(1):17–28.

Kunateh, Masahudu Ankiilu. 2009. "Tribalism and ethnicity could plunge Ghana into anarchy." *GhanaDot*. April 15.

Lindberg, Staffan and Kieth Weghorst. 2010. "Are Swing Voters Instruments of Democracy or Farmers of Clientelism? Evidence from Ghana." Working Paper.

Lindberg, Staffan and Minion K.C. Morrison. 2008. "Are African Voters Really Ethnic or Clientalistic? Survey Evidence from Ghana." *Political Science Quarterly* 123(1):95–122.

Long, James and Barak Hoffman. forthcoming. "Party Attributes, Performance and Voting in Africa." *Comparative Politics*.

Long, James and Clark Gibson. 2009. "The Presidential and Parliamentary Elections in Kenya, December 2007: Evidence from an Exit Poll." *Electoral Studies* 28(3).

McCrummen, Stephanie. 2007. "Tribalism Isn't on the Ballot, But in Kenya It's Key Issue; Presidential Election Is Most Competitive Since Independence." *The Washington Post*. December 26.

Mulowa, Andrew. 2011. "Don't Vote On Tribal Lines, Sata Urges Magoye." *Times of Zambia*. November 24.

Odunze, Hamilton. 2013. "Tribalism a barrier to progressive democratic dialogue." *The Vanguard*. March 04.

Posner, Daniel. 2004*a*. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48(11):849–863.

Posner, Daniel. 2004*b*. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi." *American Political Science Review* 98(4).

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

Posner, Daniel and David Simon. 2002. "Economic Conditions and Incumbent Support in Africa's New Democracies: Evidence from Zambia." *Comparative Political Studies* 35.

Reeves, Keith. 1997. *Voting Hopes or Fears?: White Voters, Black Candidates, and Racial Politics in America*. New York: Oxford University Press.

Robinson, Amanda Lea. 2009. "National versus Ethnic Identity in Africa: State, Group and Individual-Level Correlates of National Identification." Afrobarometer Working Paper 22.

Rubongoya, Joshua B. 2007. *Regime Hegemony in Museveni's Uganda*. New York: Palgrave Macmillan.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Walton, Hanes and Robert Charles Smith. 2000. *American politics and the African American Quest for Universal Freedom*. New York: Addison Wesley Longman.

Wantchekon, Leonard. 2003. "Clientelism and Voting Behavior: A Field Experiment in Benin." *World Politics* 55(3):399–422.

Weghorst, Keith. 2011. "Political Attitudes and Response Bias in Semi-Democratic Regimes: A Survey Experiment Comparing the List Experiment and Randomized Response in Tanzania." Working Paper.

Young, Daniel. 2009. "Support You Can Count On? Ethnicity, Partisanship, and Retrospective Voting in Africa." Afrobarometer Working Paper 115.

**Appendix 1: Proctor Script**

[Greeting]. My name is [name]. In this part of the study, I will read you some descriptions of candidates who might run for president. In each round, you will hear about two candidates. I will read each description twice so that you can be sure to hear everything. Then you will fill out a secret ballot to indicate which of the two candidates you would prefer. If you prefer the first candidate, you will mark the circle. If you prefer the second candidate you will mark the square. Once you are finished marking your ballot, you will drop it in the box up here. Then you will sit down [If voting aloud, add and tell everyone how you voted] and then I will read you two more candidates. You will vote a total of three times.

Okay, does everyone have their ballots and a pen? Please find the ballot with one [or two or three as appropriate ] small shape(s) in the corner here. [Indicate where on ballot to look. Check to make sure that everyone is using the ballot that corresponds with the voting round.] Now, I am about to read you the descriptions of two candidates - if you like the first one, tick the circle. If you prefer the second candidate, tick the square. Do not show anyone what you are marking. When you are finished, fold the ballot and put it in this box.

**Appendix 2: Hometown Cues**

The hometowns of the candidates in the voting simulation were assigned randomly from the following table with weight on the hometowns representing the locally dominant ethnic groups (usually those hometowns closest to the survey site). The hometowns in the list were chosen in cooperation with local enumerators with the criteria that they a) clearly signal a particular ethnicity and b) be familiar to respondents from all parts of the country. Though the particular location chosen to signal an ethnicity was not always the location most heavily populated by members of that particular ethnicity – often because it was a city and therefore somewhat diverse – it was chosen because it was the most recognizable location within the ethnicity's traditional home area. When a survey site was populated with respondents from a group not on the list, as with the tiny Bwamba group in Bundibugyo, coethnicity was signaled using the location of the survey site.

The hometown that would signal that a candidate was a Muganda was the most difficult to select. The seat of Buganda Kingdom is in Kampala, but as Uganda's capital city, Kampala is far too diverse to clearly signal a particular ethnicity. On my enumerators' recommendation, I selected Mukono, which is the second-largest city in Buganda Kingdom. In addition, enumerators disagreed about how to signal that a candidate was a Musoga; as shown in the table, two towns were used and candidates from either town were coded as coethnic to Basoga respondents. Results are robust to either or both being coded as coethnic.

**Appendix 3: Hometown Cue for Each Ethnicity in the Sample**

| Desired Ethnic Cue | Candidate Hometown | Region | Ethnic Pop. in Town | Ethnic Pop. in Country | Note |
|---|---|---|---|---|---|
| Acholi | Gulu | Northern | 91% | 5% | |
| Alur | Nebbi | Northern | 73% | 2% | |
| Itesot | Soroti | Eastern | 82% | 6% | |
| Japadhola | Tororo | Eastern | 41% | 1% | Parishes in Tororo up to 100% Japadhola |
| Kakwa | Koboko | Northern | 71% | 1% | |
| Karamajong | Moroto | Northern | 86% | 1% | |
| Kumam | Kaberamaido | Northern | 75% | 1% | |
| Lugbara | Arua | Northern | 77% | 4% | |
| Madi | Moyo | Northern | 54% | 1% | Parishes in Moyo up to 90% Madi |
| Mufumbira | Kisoro | Western | 87% | 2% | |
| Muganda | Mukono | Central | 72% | 17% | |
| Mugishu | Mbale | Eastern | 99% | 5% | |
| Muhororo | Rukungiri | Western | 48% | 1% | Parishes in Rukungiri up to 99% Bahororo |
| Mukiga | Kabale | Western | 96% | 7% | |
| Mukonzho | Kasese | Western | 85% | 2% | |
| Munyankole | Mbarara | Western | 73% | 10% | Mbarara is seat of Ankole Kingdom |
| Munyoro | Hoima | Western | 55% | 3% | Hoima is seat of Bunyoro Kingdom |
| Musoga | Iganga | Eastern | 90% | 8% | |
| Musoga | Jinja | Eastern | 69% | 8% | Jinja is seat of Busoga Kingdom |
| Mutooro | Fort Portal | Western | 76% | 2% | Fort Portal is seat of Tooro Kingdom |
| (Mu)Samia | Busia | Eastern | 45% | 1% | Parishes in Busia District up to 95% Samia |

**Appendix 4: Candidate Characteristics**

| Attribute | Possible options |
|---|---|
| Ethnicity | One of 21 hometowns |
| Education | Bachelor's degree in Tourism, Horticulture, Forestry or Theater<br>Master's degree in Business, Law, Economics, or African Development |
| Prior Office | None (businessman)<br>Village chief, county chief, or MP |
| Record | None (sentence omitted)<br>Paved road, created jobs, built clinic |
| Platform | Improve schools, improve health care, create better jobs |

**Appendix 5: Model Predicting Support for the Incumbent**

$$\widehat{p_{inc}} = a + b_1 Dem + b_2 Socio + b_3 Pocket + b_4 Approval + b_5 Govt \quad (2)$$

where Dem is a vector of demographic chatacteristics, Socio and Pocket are sociotropic and pocketbook economic evaluations, respectively; Approval is presidential approval rating and Govt takes a one if the respondent believes the government is running the survey.

**Appendix 6: Model Predicting Response**

$$\widehat{p_{report}} = a + b_1 Dem + b_2 News + b_3 Free + b_4 Coop + b_5 Govt + b_6 Enum \quad (3)$$

where Dem is a vector of demographic chatacteristics, New is a measure of news listenership, Free takes a one if the respondent believes he is safe to say what he thinks, Coop is the enumerator's report of how cooperative the respondent was, Govt takes a one if the respondent the government is running the survey and Enum takes a one if the respondent and enumerator are of the same ethnicity.

*AFROBAROMETER WORKING PAPERS*

No. 143 Ali, Merima, Odd-Helge Fjeldstad and Ingrid Hoem Sjursen. "To Pay or Not to Pay? Citizens' Attitudes Towards Taxation in Kenya, Tanzania, Uganda and South Africa" 2013

No. 142 Bodenstein, Thilo. "Ethnicity and Individual Attitudes Towards International Investors: Survey Evidence from Sub-Saharan Africa" 2013

No. 141 Bandyopadhyay, Sanghamitra and Elliott Green. "Pre-Colonial Political Centralization and Contemporary Development in Uganda" 2012

No. 140 Sacks, Audrey. "Can Donors and Non-State Actors Undermine Citizens' Legitimating Beliefs" 2012

No. 139 Justesen, Mogens K. and Christian Bjørnskov. "Exploiting the Poor: Bureaucratic Corruption and Poverty in Africa" 2012

No. 138 De Luca, Giacomo and Marijke Verpoorten. "From Vice to Virtue? Civil War and Social Capital in Uganda" 2012

No. 137 Kerr, Nicholas. "Perceptions versus Reality: Assessing Popular Evaluations of Electoral Quality in Africa" 2011.

No. 136 Resnick, Danielle and Daniela Casale. "Political Participation of Africa's Youth: Turnout, Partisanship, and Protest" 2011.

No. 135 Conroy-Krutz, Jeffrey and Carolyn Logan. "Museveni and the 2011 Ugandan Election: Did the Money Matter?" 2011.

No. 134 Carter, Danielle. "Sources of State Legitimacy in Contemporary South Africa: A Theory of Political Goods" 2011.

No.133 Justesen, Mogens K. "Too Poor to Care? The Salience of Aids in Africa" 2011.

No.132 Kasara, Kimuli. "Separate and Suspicious: Local Social and Political Context and Ethnic Tolerance in Kenya" 2011.

No.131 Mattes, Robert. "The 'Born Frees': The Prospects for Generational Change in Post-Apartheid South Africa" 2011.

No.130 Baldwin, Kate. "When Politicians Cede Control of Resources: Land, Chiefs and Coalition-Building in Africa" 2011.

No.129 Conroy-Krutz, Jeffrey and Dominique Lewis. "Mapping Ideologies in African Landscapes" 2011.

No.128 Logan, Carolyn. "The Roots of Resilience: Exploring Popular Support for African Traditional Authorities" 2011.

No.127 Bratton, Michael, Ravi Bhavnani and Tse-Hsin Chen. "Voting Intentions in Africa: Ethnic, Economic or Partisan?" 2011.

No.126 D'Arcy, Michelle. "Why Do Citizens Assent to Pay Tax? Legitimacy, Taxation and the African State" 2011.

No.125 LeBas, Adrienne. "Ethnicity and the Willingness to Sanction Violent Politicians: Evidence from Kenya" 2010.

No.124 Mattes, Robert, Francis Kibirige and Robert Sentamu. "Understanding Citizens Attitudes to Democracy in Uganda" 2010.

No.123 Logan, Carolyn and Robert Mattes. "Democratizing the Measurement of Democratic Quality: Public Attitude Data and the Evaluation" 2010.