NBER WORKING PAPER SERIES

HOW STRONG ARE ETHNIC PREFERENCES?

Lars Ivar Oppedal Berge
Kjetil Bjorvatn
Simon Galle
Edward Miguel
Daniel N. Posner
Bertil Tungodden
Kelly Zhang

Working Paper 21715
http://www.nber.org/papers/w21715

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2015

We thank Ana Beatriz Aguilar, Felipe Gonzalez, Ranveig Falch, Ruth Carlitz, and Dounia Saeme for excellent research assistance, and the Busara Center for Behavioral Economics, the Economics and Social Research Foundation (ESRF), and The Choice Lab for administrative support. We are grateful to Obie Porteus, Gwyneth McClendon, and Jessica Leight for sharing data. We are grateful for funding from the Governance Initiative at JPAL-MIT, the Peder Saether Foundation, and the Norwegian Research Council. Galle acknowledges support from U.C. Berkeley IBER, Zhang acknowledges support from the NSF Graduate Research Fellowship Grant No. DGE-114747. We have received extremely useful comments from Michal Bauer, Julie Chytilova, Guy Grossman, Johannes Haushofer, Jonas Hjort, Dean Karlan, Steve Krasner, David Laitin, Rohini Pande, Amanda Robinson, Raul Sanchez de la Sierra, and seminar audiences at the Norwegian School of Economics, U.C. Berkeley, Caltech, Harvard, UCLA, Stanford, Columbia, University of Washington, Princeton, Yale, U.C. Merced, Stockholm School of Economics, the Annual World Bank Conference on Africa, the Barcelona GSE Summer Forum, and CEGA. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How Strong are Ethnic Preferences?
Lars Ivar Oppedal Berge, Kjetil Bjorvatn, Simon Galle, Edward Miguel, Daniel N. Posner, Bertil Tungodden, and Kelly Zhang
NBER Working Paper No. 21715
November 2015, Revised April 2016
JEL No. C90,H41,O43

ABSTRACT

Ethnic divisions have been shown to adversely affect economic performance and political stability, especially in Africa, but the underlying reasons remain contested, with multiple mechanisms potentially playing a role. We utilize lab experiments to isolate the role of one such mechanism—ethnic preferences—which has been central in both theory and in the conventional wisdom about the impact of ethnic differences. We employ an unusually rich research design, collecting multiple rounds of experimental data with a large sample of 1,300 subjects in Nairobi; employing within-lab priming conditions; and utilizing both standard and novel experimental measures, as well as implicit association tests. The econometric approach was pre-specified in a registered pre-analysis plan. Most of our tests yield no evidence of differential altruism towards coethnics relative to non-coethnics. The results run strongly against the common presumption of extensive ethnic bias among ordinary Kenyans, and suggest that other mechanisms may be more important in explaining the negative association between ethnic diversity and economic and political outcomes.

Lars Ivar Oppedal Berge
Norwegian School of Economics
Bergen, Norway
Lars.Ivar.Berge@nhh.no

Kjetil Bjorvatn
Norwegian School of Economics
Bergen, Norway
Kjetil.Bjorvatn@nhh.no

Simon Galle
University of California
Berkeley, CA 94720 USA
sgalle@econ.berkeley.edu

Edward Miguel
Department of Economics
University of California, Berkeley
530 Evans Hall #3880
Berkeley, CA 94720
and NBER
emiguel@econ.berkeley.edu

Daniel N. Posner
UCLA
Los Angeles, CA USA
dposner@polisci.ucla.edu

Bertil Tungodden
Norwegian School of Economics
Bergen, Norway
Bertil.Tungodden@nhh.no

Kelly Zhang
Stanford University
Palo Alto, CA USA
kwzhang@stanford.edu

A data appendix is available at http://www.nber.org/data-appendix/w21715
A randomized controlled trials registry entry is available at https://www.socialscienceregistry.org/trials/16

1. Introduction

Ethnic divisions have been shown to adversely affect economic performance and political stability, especially in Africa (Easterly and Levine 1997, Cederman et al. 2007, 2011), but the underlying mechanisms remain contested. One channel that has played a particularly central role in theory and is at the root of the conventional wisdom about why ethnicity matters is that individuals exhibit greater altruism for coethnics, and even antipathy toward non-coethnics (Vigdor 2002). Explanations of this sort, which focus on *ethnic preferences*, contrast with explanations that stress the *institutional and technological factors* that raise the transaction costs associated with cross-group interactions, including difficulties in communication, imposing social sanctions and enforcing social norms across group lines (Hardin 1995; Miguel and Gugerty 2005; Habyarimana et al. 2007).

We study the strength of ethnic preferences in Nairobi, Kenya, a setting with well-documented and politically salient ethnic divisions, utilizing lab experiments to isolate this mechanism from other factors. While the term "ethnic preferences" is sometimes also used to refer to the correlation between group membership and preferences over public policies (e.g., Alesina, Baqir and Easterly 1999, Alesina and LaFerrara 2005, Lieberman and McClendon 2013), we use it here to refer specifically to the different levels of altruism in within- versus cross-group interactions, a phenomenon we also refer to as "ethnic bias". Our approach is motivated by the frequency, in both academic scholarship and popular accounts, with which outcomes in Africa are linked to ethnicity by invoking this logic.

We employ an unusually rich research design that involves multiple rounds of experimental data with a large sample of over 1300 subjects, including one round immediately prior to a national election, when conventional wisdom would suggest that coethnic bias should be particularly strong (Eifert et al. 2010, Snyder 2000). We

supplement this variation in real-world timing with within-lab priming designed to increase the situational salience of particular issues and dimensions of social identity. We measure preferences using both standard experimental games (e.g., Dictator and Public-good games) and a more novel lab activity (the Choose-Your-Dictator game, which captures expectations about the altruism of others). As a further check, we employ an implicit association test (IAT) to capture underlying preferences free from experimenter demand effects, a potential concern in contexts where overt expression of ethnic bias may be socially undesirable. The design and econometric approach were pre-specified in a registered pre-analysis plan, incorporating adjustments for multiple hypothesis testing. Pre-analysis plans are increasingly used in field experiments (Casey et al. 2012) but their use to date in laboratory experimental studies has been limited.

Given the widely held views on both the strength and the negative implications of ethnic preferences, our results are as striking as they are optimistic. Most of our tests yield no evidence of coethnic bias, while a few show evidence of a small degree of bias. This lack of meaningful ethnic bias holds across multiple experimental measures and well-powered statistical tests, including the IAT. Figure 1 summarizes evidence from the Dictator game and the Public-good game. In both cases, the average level of contribution to coethnic and non-coethnic partners (where identification of partner ethnicity is introduced in a way we discuss below) is strikingly similar, within one percentage point: 35.6% versus 35.4% in the Dictator game (Panel A), and 46.2% versus 46.4% in the Public-good game (Panel C). Participants' beliefs about the contributions of others in these games are similarly unaffected by their ethnic backgrounds, at 49.3% to 48.4% (Panel B) and 53.9% to 53.1% (Panel D), respectively. None of these differences are statistically distinguishable from zero. In the IAT, the average bias against members of other ethnic groups is just 0.079 standard deviation units, roughly one sixth of the average bias demonstrated by U.S. whites against blacks (Nosek et al. 2007).

These precisely estimated null effects are robust: they hold among almost all of the demographic subgroups we pre-specified (including by gender, ethnic group, and education); in the experimental round close in time to national elections, as much as in the round conducted in the previous year; and across a range of priming conditions, including primes for ethnic identity, political competition, and national identity. The core null finding is also not the result of selective presentation of results on our part: the econometric approach was pre-specified and, in a novel test, we present the distribution of statistical significance levels for results contained in the main tables versus the full set of pre-specified results.

It is a testament to the broad acceptance of the ethnic preferences mechanism that many readers will find these results surprising—especially given our Kenyan research site, which is commonly associated with ethnic rivalry and which witnessed ethnic violence in the aftermath of the disputed 2007 national elections that led to more than a thousand deaths and the displacement of hundreds of thousands of people. Moreover, modernization theorists have argued that African urban environments like Nairobi are especially prone to ethnic antagonisms due to the combination of social heterogeneity and the heightened competition for jobs and resources (Bates 1983, Melson and Wolpe 1970). So vis-a-vis both history and theory, our findings are unexpected.

It is worth being absolutely clear that we are not downplaying the role of ethnicity in the study setting: ethnic divisions remain a prominent feature of contemporary Kenyan society. Our findings simply suggest that the salience of ethnicity may be due to mechanisms other than ethnic preferences. Indeed, notwithstanding the conventional wisdom, there is actually more empirical evidence in favor of other channels than there is for explanations rooted in ethnic preferences, at least in African cases. Miguel and Gugerty (2005) argue that all ethnic groups in western Kenya have strong preferences for funding local schools, but that diverse communities have far worse voluntary local fundraising outcomes due to their inability to sanction free-

riders. Building on the seminal work of Barkan and Chege (1989), Burgess et al. (2015) document large-scale distortions in public roads investment in Kenya favoring the president's ethnic group. In their model, this is an equilibrium choice due to the instrumental political benefits for rulers, but it does not rely on any coethnic bias (although they cannot rule out that such bias is playing some role). Habyarimana et al. (2007) provide further evidence from lab experiments in urban Uganda that ethnic preference explanations are less powerful than accounts emphasizing the different norms governing coethnic and non-coethnic interactions, and the role of within-group sanctions to enforce them.

At the same time, several other studies find mixed evidence on coethnic preferences in African settings (Carlson 2015, Michelitch 2015, Dionne 2014, Grossman and Honig 2015, Hjort 2014, Marx et al 2015, Loewen, Eady and Fowler 2014, Jeon 2013, Voors et al. 2012).[1] The current study contributes to this emerging literature by employing larger samples (and thus greater statistical power) and a larger number of experimental measures and tests, all using a pre-specified set of research hypotheses. We believe the results thus serve to crystallize that coethnic bias in African societies is often less pronounced than is widely believed in either scholarly or popular debates.

The rest of the paper is organized as follows. Section 2 presents a theoretical framework relevant for disentangling ethnically biased behavior and ethnic preferences in the context of the lab games; section 3 describes the setting, experimental design, and econometric approach; section 4 describes the data and empirical results; and the final section concludes. The supplementary appendix contains additional theoretical results (appendix A), tables and figures (appendix B), the lab experimental protocol and survey

[1] Michelitch (2015) investigates whether proximity to Ghana's 2008 election affects how coethnics and non-coethnics (as well as copartisans and non-copartisans) bargain over taxi fares, and finds evidence consistent with coethnic bias. While her research is similar in exploiting variation in real world context, her field experimental design does not focus on isolating coethnic preferences.

instruments (appendix C), pre-analysis plans (appendix D), and the full set of pre-specified results (appendix E).

2. A model of ethnic preferences and behavior

In the empirical analysis, we employ lab games to measure individual-level ethnic biases in behavior. The theoretical framework developed below describes what can be learned about ethnic bias in preferences from observing individual-level play in the Dictator and Public-good games. Specifically, it reveals that, under the weak assumption that subjects are not perfectly altruistic, a lack of observed coethnic bias in behavior implies that there is no ethnic bias in individual preferences. Critically, it also shows how ethnic bias in behavior is not necessarily driven by bias in *ethnic preferences*.

The model builds on Cappelen et al. (2007, 2013), where individuals make a trade-off between selfish and other-regarding concerns when sharing a sum of money. We extend the original model by introducing ethnic preferences so that it matches the versions of our lab games where subjects have some information about the likely ethnicity of their partners. We present the baseline model in the context of a Dictator game; afterward, we discuss how the model relates to the Public-good game.

Let m_i be the individual's fairness ideal. This could be an even split of the endowment, although the exact value does not matter for measuring the degree of observed ethnic bias in behavior. We introduce an ethnic bias in preferences, given by q_i, and an ethnicity indicator variable n_j, which takes the value one if the counterpart is a non-coethnic and zero if a coethnic. The utility of an individual i is:

$$u_i(t) = (1 - t_i) - \frac{(t_i - m_i)^2}{2b_i(1 + n_j q_i)} \quad (1)$$

where the endowment of money is given by 1, t_i is the transfer to the other player, and b_i is the degree of egoism ($b_i \geq 0$), i.e., $b_i = 0$ denotes perfect altruism.

The first term in equation 1 captures the utility component based on what the decision-maker keeps for himself, and the second term is the decision-maker's disutility due to deviations from the fairness norm. Note that the disutility from deviations from this norm depends on three factors, b_i, q_i and n_j. In the limit as $b_i \to \infty$, the second term tends to zero as egoism increases, and a fully egoistic decision-maker chooses $t_i = 0$. At the other end of the spectrum, for fully altruistic individuals $b_i \to 0$, the second term receives all the weight in the decision process, and $t_i = m_i$. The degree of ethnic bias in preferences q_i has an analogous effect when the other party is a non-coethnic, i.e., for $n_j = 1$. In that case, the larger is q_i, the lower is the weight given to the fairness term, and the smaller is the transfer to non-coethnics.

Note that for $b_i(1 + n_j q_i) \geq m_i$ we have a corner solution, where the decision maker keeps all the money for herself, $t_i^* = 0$. However, for an interior solution, the first order condition for sharing implies:

$$t_i^* = m_i - b_i(1 + n_j q_i) \tag{2}$$

Sharing with a coethnic (CE) is then given by:

$$t_i^{CE} = m_i - b_i \tag{3a}$$

and with a non-coethnic (NCE):

$$t_i^{NCE} = m_i - b_i(1 + q_i) \tag{3b}$$

A low level of sharing in this framework can be due to three factors: (i) a fairness ideal that suggests a low level of sharing; (ii) a high level of egoism; or, when interacting with a non-coethnic, (iii) a high level of ethnic bias in preferences.[2]

The ethnic bias in behavior in the dictator game is thus given by:

$$B^D = t_i^{CE} - t_i^{NCE} = b_i q_i \tag{4}$$

[2] An additional theoretical possibility that may be difficult to distinguish from ethnic bias in preferences is the existence of different sharing norms (m_i) when facing a coethnic and a non-coethnic. As shown below, there is no indication of significantly different contributions to coethnics and non-coethnics in our data, providing little evidence for either ethnic bias in preferences or different sharing norms.

that is, how much more the decision-maker contributes when facing a coethnic instead of a non-coethnic. Note that the value of m_i is differenced out and does not matter for our expression of ethnic bias.

A key implication of the model is that a strong observed ethnic bias in behavior ($b_i q_i$) is not necessarily the result of a high ethnic bias in preferences alone (high q_i), as it may be driven primarily by egoism (high b_i). Another important implication of the model for our results is that, in a population with imperfect altruism ($b_i > 0$) – a realistic pattern that characterizes nearly all populations that have played the Dictator game – a finding of no ethnic bias in behavior ($B^D = b_i q_i = 0$) implies no ethnic bias in preferences ($q_i = 0$). This implication of the model will allow us to draw conclusions about the extent of ethnic preferences in our sample based on ethnic bias in behavior.

So far, we have analyzed the implications of our model in the context of the Dictator game. Similar reasoning applies to the Public-good game, where we can employ an analogous utility-structure for respondent i when making a contribution t_i to the group fund. The main difference lies in the fairness-norm becoming group-dependent. Specifically, we assume that the fairness term is now given by $m_i^g = f\left(E_{g,-i}[t]\right)$, where $E_{g,-i}[t]$ is individual i's expectation about other group-member contributions to the group fund of group g, and we assume $f' > 0$. As such, the difference in contributions in a coethnic versus non-coethnic group reflects both an ethnic bias in preferences (as above), and in expectations about others' contributions. If other group members are expected to contribute less in a non-coethnic setting compared to a coethnic setting, this should amplify the observed ethnic bias in terms of Public-good game contributions.

In the lab, we also collected data on respondents' stated beliefs about the amounts other group members will contribute. Since cooperation can unravel in voluntary Public-good games, as individuals tend to have a desire to contribute less

than they think that others are contributing (Fischbacher and Gachter 2010), we also examine the difference between individual i's contribution t_i and her expectations about others' contributions, m_i^g. This difference is often interpreted as capturing the degree of free-riding. Under further assumptions (discussed in Appendix A), the difference in contributions minus beliefs in a coethnic versus a non-coethnic Public-good game setting is again $b_i q_i$, as in the Dictator game, although the interpretation of this measure is slightly different: in the Dictator game, it captures ethnic bias in generosity, whereas in the Public-good game it can be thought of as ethnic bias in free-riding. As above for the Dictator game, in a population with non-zero egoism ($b_i > 0$), a finding of no ethnic bias in behavior in the Public-good game implies no ethnic bias in preferences ($q_i = 0$).

3. Setting, experimental protocol, and research design

3.1. Setting, sample and timing

The study sample draws from two low-income neighborhoods in Nairobi, Kenya, namely, Kibera, which was a focal point for the 2007-08 post-election violence, and Viwandani. Subject recruitment was conducted through, and the experiments were carried out at, the Busara Center for Behavioral Economics, a high-quality lab setting similar to those in U.S. and European research institutions, with touch screen monitors and headphones for audio instructions (to enable illiterate people to participate). To enhance representativeness, recruitment was stratified by ethnic group, with five of the six largest ethnic groups in Kenya well-represented, namely, the Kikuyu, Luo, Luhya, Kamba, and Kisii groups (listed in order of size in both the Kenyan population and our study sample); subject recruitment only occurred among members of these five groups. Taken together, these five groups make up 82% of the Nairobi population according to

1999 Kenyan census data.[3] They also contain the two groups with the deepest historical rivalry, Kikuyu and Luo. (The only major Kenyan ethnic group not represented in the sample is the Kalenjin, whose members do not reside in large numbers in either of the two recruitment neighborhoods.)

The Kikuyu and Luo ethnic rivalry has played an important role in Kenyan politics for decades. After independence in the 1960's, national politics was dominated by the rivalry between Oginga Odinga (an ethnic Luo) and Jomo Kenyatta (a Kikuyu), the country's first president. The 2007 post-election violence ended once the two presidential frontrunners Raila Odinga (a Luo and son of Oginga) and Mwai Kibaki (an ethnic Kikuyu) reached a power-sharing agreement. In 2013, the sons of the two post-independence leaders were the frontrunners, with Raila Odinga losing to Uhuru Kenyatta, the current president of Kenya.

There were two data collection rounds in our study, distinguished by their proximity to the March 2013 presidential and parliamentary elections. We call the first round from July to August 2012 the *Non-Election round*, and the second round from January to February 2013 the *Election round*. The macroeconomic environment was relatively stable during the study period, as captured in major commodity prices during 2011 to 2013 (see Appendix Figure S1). The total sample size is 1,362 distinct individuals, with 608 individuals in the Non-Election round and 754 in the Election round. Participants were recruited in each round following an equivalent protocol.

Two lab sessions were held each day, each lasting two to three hours and including up to 20 participants. As compensation, individuals were reimbursed for transport and given a "show-up fee" (in addition to their payouts from the games played). Payments took place by mobile money transfer after the session had ended. All games were paid out, and results of each game were not revealed until after the session.

[3] These five groups make up roughly sixty percent of the national population, according to 2009 Kenyan census data from the Kenya National Bureau of Statistics (KNBS).

In order to minimize the probability that participants discovered each other's ethnicity, they were asked not to talk to each other and they sat in separate cubicles once they entered the lab premises. A discussion of the experimental lab protocol, including details of recruitment, payment, and lab procedures is contained in Appendix C.

Figure 2 illustrates the structure of the lab activities. Each lab session was divided into two parts, both chronologically and conceptually, featuring a part with standard anonymous games followed by a part with "profiled" games, in which individuals were provided some information about their partners.

In the standard part, participants were randomly paired with individuals about whom they were provided no information. These partners were individuals who had played in pilot lab rounds and were randomly chosen to receive payouts based on the play of the games. Although participants had no way of knowing the specific ethnicities of the partners they were paired with, they knew that they were recruited from Nairobi's multiethnic population. Hence, if they had reason to wonder about their partners' ethnic backgrounds, participants could reasonably infer that they were quite unlikely to be coethnics; as we discuss below, the largest ethnic group in our sample (Kikuyus) comprised just 32% of participants (Table 1).

In the profiled part of the lab, participants were paired with individuals about whom they received three pieces of background information: their education, age, and home region. The objective was to provide information with which a participant could make an inference about her/his partner's ethnic affiliation without being completely overt about the focus of the experiment. The inclusion of information about the partner's age and education was meant to obscure the ethnic focus of the study; as we discuss below, it appears to have been largely successful. For instance, one actual profile was presented as follows: "*This participant is originally from near Kisumu, went to school up to Form 2 and was born in 1979.*" Home region – here Kisumu, a city in the western part of the country with an overwhelming Luo ethnic majority (84% in the 1999

national census) – is a strong indicator the participant's ethnic background. Local interviews have confirmed that the vast majority of Kenyans would conclude correctly that this partner was an ethnic Luo. The order of the three pieces of information (home region, school completion, and birth year) was varied across subjects so as to avoid bias stemming from order effects.

The profiles of the other players with whom participants were paired were randomly assigned. As in the standard games, the profiles were drawn from among the individuals who participated in the pilot sessions, only this time we limited the partner sample to those whose home regions provided unambiguous clues about their ethnicity. To ensure a sufficiently large number of coethnic matches, we further limited the partner sample to individuals from the two largest ethnic groups, the Kikuyu and the Luo, groups whose high political relevance insures that many of the interactions we study were between members of groups whose relations have been historically fraught.

As laid out in our pre-analysis plan (see appendix D), and in accordance with the ethnic political coalitions that emerged in advance of the 2013 national election, we choose to categorize Luos together with Luhyas and Kisiis (two other western Kenyan groups long in coalition with Luos in national politics) as coethnics in the analysis, as distinct from the larger Kikuyus. To be conservative, we pre-specified that the Kamba would be excluded from the main analysis due to the shifting political alliances of their leaders during 2012 and 2013 (but the results do not differ if they are included and grouped with the Kikuyu, the group that is culturally and geographically closest to them). The main empirical results are also unchanged if the sample is restricted to just Kikuyu and Luo subjects alone, who together make up roughly half the sample, ignoring allied ethnic groups, as we discuss below.

Each lab session consisted of three main games: the Dictator Game (DG), the Public-good game (PG), and the Choose-Your-Dictator (CYD) game. Within the profiled parts of the DG and PG, the order of the coethnic and non-coethnic games was

randomized. The CYD, by contrast, was always played with at least some partner information, although there are two information variants to this game, described below.

In addition to the behavioral games, participants were exposed to a randomly assigned priming treatment designed to increase the situational salience of (1) ethnic-cultural differences, (2) ethnic-political differences, (3) political competition, or (4) national identity. There was also a control group in which participants were exposed to a neutral prime. We hypothesize that the priming treatments might affect the degree of ethnic preferences (q_i) or egoism (b_i), and thus affect coethnic bias in behavior.

The primes were designed to serve as experimental counterparts to the real world variation in proximity to elections that we examine and, in the case of the national identity prime, to explore the effects on ethnically motivated behavior of invoking a common superordinate identity (Kramer and Brewer 1984, Gaertner and Dovidio 2000). The political competition prime serves to provide experimental variation in "exposure" to an election period. As such, we might expect this priming treatment to identify the same mechanisms as we observe when comparing the 2012 Non-election round data to the 2013 Election round. To the extent that ethnicity and politics are tightly linked in subjects' minds, priming them to ethnic identity may have a similar impact. It may also have an effect independent of its political connections, simply by alerting subjects to the salience of ethnic differences. The two types of ethnic primes are meant to mimic the daily exposure that individuals have to appeals to shared ethnic culture (the ethnic-cultural prime), and to blatant tribal politics (the ethnic-political prime). The latter may operate by increasing perceptions of intergroup competition, which may exacerbate individual level discrimination along group lines (Bowles and Gintis 2004; Shayo 2009). We clearly are not mimicking (or attempting to mimic) hate speech with the primes, but do hope to capture a similar degree of situational salience as one would experience hearing bits of a politician's speech or radio discussion.

To reduce the likelihood that priming effects would dissipate, priming was implemented at four distinct points within the lab session (see Figure 2). The primes were administered between each set of games as short, three question on-screen quizzes. As an illustration, one of the political competition priming questions asked: "*How many political candidates are running for the Presidency?*" One of the ethnic-cultural priming questions was: "*This greeting comes from which region: 'Orie'?*" Questions specific to the priming conditions were inter-mixed with neutral questions, mainly focused on Kenyan popular culture or everyday life in Nairobi, for example: "*How often do you ride a matatu (mini-bus) every week?*" Each quiz included two priming questions and one neutral question. Those in the control prime group were asked only neutral questions. See Appendix C for the full set of priming questions.

We implemented a 16-piece Raven Progressive Matrices test at the start of each lab session for a measure of individual cognitive ability (Raven 2008). We use these scores both to confirm balance across treatment groups in this potentially important characteristic and as a control in our analyses. We normalize the Raven's test score in our sample, resulting in a measure with mean zero and standard deviation of one.

At the end of the session, subjects were administered an IAT, and then asked survey questions to capture self-reported political preferences and attitudes. The questions gathered information on subjects' preferences among the presidential candidates in the upcoming election, their actual vote choice in the previous election, their attitudes regarding the 2007 post-election violence, whether they had received items or cash from political campaigns, and the number of political rallies they had attended. We also asked the standard Afrobarometer question on the strength of respondents' ethnic attachment relative to other identity dimensions (Eifert et al. 2010).

3.2. Pre-analysis plans

We pre-specified our analytical approach in a pre-analysis plan registered in the American Economic Association Social Science Registry (ID# AEARCTR-0000016). The analysis described below closely follows the plan (see Appendix D). The past few years have seen a rise in the use of pre-analysis plans for field experiments (Miguel et al. 2014), but this study is to our knowledge among the first to apply this approach to a laboratory experiment. A registered PAP helps address concerns of data-mining and publication bias. These concerns are especially salient in our study given its complex structure, with multiple games, primes, subsamples, and lab rounds, which allow for a plethora of potential hypothesis tests.

There are several noteworthy aspects of our application of a pre-analysis plan in this context. First, in direct response to critics who claim that pre-specification undermines flexibility, we demonstrate that it is possible to reconcile the natural evolution of research objectives over the course of a study with a commitment to pre-specification. After analyzing the data from the 2012 Non-election round, we registered a second PAP in anticipation of analyzing the Election round data, and it featured several modifications, including an additional priming condition – namely, the ethnic-political prime, which more explicitly attempted to trigger the salience of ethnic politics – as well as several econometric modifications. This approach of updating a pre-analysis plan can be generalized to any study setting where new data becomes available over time (for instance, in the case of ongoing longitudinal data collection) as long as the registration of the plan precedes access to and analysis of the new data.

Second, we present a number of analyses that were not pre-specified, clearly highlighting them for the reader. Important insights inevitably occur over the course of a study, and not every piece of analysis can be predicted in advance, making it important to maintain flexibility as long as there is full transparency about what was and was not pre-specified. A reader preferring to disregard the non-prespecified

hypotheses is free to do so, but they are made available to others who may find them useful. For instance, our discussion below of the differences in behavior between recent migrants to Nairobi versus long-term residents was not prespecified but provides an informative test of a key observable implication of one interpretation of our results.

Finally, as a check against the possibility that we are selectively emphasizing a non-representative subset of results from among the universe of all tests in the PAP (as provided in Appendix E), we also compare the distribution of p-values contained in the main tables of this article with the distribution of p-values for all hypothesis tests described in the plan. As there are hundreds of hypothesis tests described in the PAP, it is not feasible to discuss all of them in a research paper of standard length. This comparison allows us to assess whether we are disproportionately emphasizing (or de-emphasizing) statistically significant results in the main study tables.

Coffman and Niederle (2015) have recently argued against the use of pre-analysis plans for lab experiments, claiming that they are superfluous in settings where experiments can be easily replicated. While we wholeheartedly endorse Coffman and Niederle's emphasis on replication, we view replication studies and pre-analysis plans as complements rather than substitutes, and believe that our analysis, which relies on a particular context (namely, proximity to the Kenyan elections), highlights some important benefits of a pre-analysis plan. Since it is not possible to replicate our study of the 2013 Election round effects in Kenya in the absence of a time-machine, substituting replication for pre-specification is not an option. In such a context, pre-specifying our hypothesis helps to bolster the credibility of the results. The same holds for the growing number of scholars collecting lab experimental data in the context of a parallel field experiment, or where the research design is linked in some way to a real-world natural policy experiment or "shock" (Fisman et al. 2014, Michelitch 2015, among many others). In cases where lab experimental researchers only have "one shot" at a particular type of data – due to access, cost, or timing – pre-analysis plans can be a useful tool.

3.3. Dictator Game

The Dictator game aims to capture an individual's altruism towards others. Participants were given an endowment of 50 Ksh ($0.60) – equivalent to roughly an hour's wage for many Nairobi workers – of which they had to decide how much to give away to an anonymous partner with whom they were randomly paired. Subjects played both standard and profiled version of the game, with the latter played twice: once with a coethnic partner once and once with a non-coethnic partner. This allows us to measure the degree of coethnic bias in the behavior of each participant, as captured by the difference in transfers to the coethnic versus non-coethnic partner.

The standard Dictator game was played in both the Election and Non-election rounds but, due to a programming error in the lab software, only the coethnic version of the profiled game was played in the Non-election round. This implies that we can estimate the extent of ethnic bias in the Dictator game only in the Election round. (There was no such coding error for the Public-good game or the Choose-Your-Dictator game, so for those games we are able to measure coethnic bias in both the Non-election and Election rounds, as described below.)

In the analysis, the outcome of interest is the percentage of the endowment transferred to the partner. We first focus on how partner coethnicity affects transfers, and how coethnicity interacts with election proximity, and then estimate the effect of the priming treatments.

The first specification includes data from both the standard and profiled games:

$$Y_{ij} = \alpha + \beta_1 Coethnic_{ij} + \beta_2 Election_i + \beta_3 Coethnic_{ij} * Election_i + \beta_4 Noncoethnic_{ij} * Election_i + \varepsilon_{ij} \quad (5)$$

Y_{ij} is the transfer (in percent of the endowment) by subject i to their partner in game j, $Coethnic_{ij}$ ($Noncoethnic_{ij}$) is an indicator variable for being paired with a coethnic (non-

coethnic) partner in game *j*, $Election_i$ indicates whether respondent *i* is observed in the Election round, and ε_{ij} is the error term, clustered by individual. The lack of a non-coethnic profiled Dictator game in the Non-Election round explains the missing non-interacted $Noncoethnic_{ij}$ term. The estimate of coethnic bias in behavior in the Election round is given by the difference in coefficient estimates on the coethnic partner and non-coethnic terms, and the formal hypothesis test is that $(\beta_1 + \beta_3) = \beta_4$. The Election round effect captured in β_2 is potentially of independent interest. As a robustness check, we also estimate a specification with covariates including ethnicity and gender, years of education, and the normalized Raven's test score.

We estimate effects of the priming treatments in two slightly different ways. In the main analysis that we report in the text, we focus on the standard Dictator game in order to provide the cleanest possible test of priming effects. Since participants could reasonably infer that anyone they were paired with was likely to be a non-coethnic, the results of the standard Dictator game provides insight into the effects of priming on levels of altruism vis-a-vis non-coethnics. We estimate the following equation:

$$Y_{ij} = \alpha + \beta_1 Election_i + \sum_{k=1}^{3} \beta_{1+k} T_k + \sum_{k=1}^{4} \beta_{4+k} (T_k * Election_i) + \varepsilon_{ij} \qquad (6)$$

The terms T_1, T_2, T_3 and T_4 are indicator variables for the four priming conditions, namely, the political-competition prime, the ethnic-cultural prime, the national prime, and the ethnic-political prime, respectively. β_1 estimates the difference in the average level of transfers across the Election and Non-election rounds, β_2, β_3, and β_4, the average effects of the first three priming treatments in the Non-election round, and β_5, β_6 and β_7 the differential effect of these primes in the Election round. β_8 is the effect of the ethnic-political prime, which was only administered in the Election round, and thus is only included in interaction with the Election round indicator.

We also analyze the effects of priming using the profiled Dictator game. This analysis provides an arguably more direct test of the impact of priming on ethnic bias,

but has the drawback of potentially cueing subjects, through both the priming and the home district information provided in the profile, to the researchers' interest in ethnicity. Hence we put more weight on the priming analyses that employ the non-profiled games. In any case, results are similar using both approaches (Appendix E).

3.4. Public-good Game

Participants in the Public-good game were given an endowment of 60 Ksh ($0.70), and could choose how much to keep for themselves versus how much to contribute to a group fund in which contributions were multiplied by two before being equally shared among the three players in the game. The contribution level thus captures an individual's willingness to share resources to make others better off, and resembles the classical prisoner's-dilemma. Compared to the Dictator game, the Public-good game's framing in terms of a "group fund" is distinct, and the recognition that other actors are also making decisions could trigger reciprocity that might affect behavior.

Before deciding how much to contribute themselves, subjects stated how much they believed each of the other players would contribute. These were non-incentivized elicitations. Subjects received no information about the two other players in the standard part of the session, where once again the other players were randomly drawn from subjects in the pilot rounds. In the second, profiled part of the lab, they were informed about each of the two other players' years of education, age and home region, just as in the profiled Dictator game. There were two types of profiled Public-good games. The first was a "mixed" group, with one coethnic player and one non-coethnic player, and the second, the "coethnic" group, in which both players were coethnics with the subject. The order in which these were played was randomized. The difference in individual contributions to the group fund across the coethnic group and the mixed group is a measure of coethnic bias. We pre-specified both a focus on contributions, as well as on "contributions minus beliefs" in the PAP (see the discussion above). For

simplicity here, we focus on contributions alone but results are very similar for the contributions minus beliefs outcomes (see Appendix E).

The main econometric specification follows the Dictator game, where $Mixed_{ij}$ is an indicator for an ethnically mixed group, $Coethnic_{ij}$ is a homogeneous coethnic group, and Y_{ij} is the contribution of individual i in game j, and all games (standard and profiled) are pooled in the analysis:

$$Y_{ij} = \alpha + \beta_1 Coethnic_{ij} + \beta_2 Mixed_{ij} + \beta_3 Election_i + \beta_4 Coethnic_{ij} * Election_i + \beta_5 Mixed_{ij} * Election_i + \varepsilon_{ij} \quad (7)$$

We estimate the extent of ethnic bias in the Non-election round in an F-test of the hypothesis that $\beta_1 = \beta_2$. We also assess if there is a differential degree of coethnic bias in the Election round by testing if $\beta_4 = \beta_5$. As in the Dictator game, we also estimate the effects of the priming interventions in the standard Public-good game (see Appendix E).

3.5. Choose-Your-Dictator Game

The Choose-Your-Dictator game was designed to capture expectations of differential levels of altruism in coethnic and non-coethnic interactions by measuring whether subjects discriminate along ethnic lines in their choice of a "leader" responsible for allocating resources to them, a conception thus broadly linked with the process of electing public officials.[4] We operationalize this by building on the Dictator games employed in this study.

Subjects ("choosers") were presented with two randomly drawn profiles of potential "Dictators", and were then asked to choose one to be "their" Dictator in a Dictator game. One of the two drawn profiles was a coethnic, the other a non-coethnic,

[4] Adida et al (2015) develop a game similar to our "Choose-Your-Dictator" game. Their experimental protocol differs as the participants went through a face-to face speed-chatting process prior to voting for a group "leader." Hence the choice of "leader" could be affected by personal characteristics out of the researchers' control, and the impact of ethnicity may be more difficult to discern.

and as before these were the profiles of actual individuals who played in the pilot rounds. As in the profiled Dictator and Public-good games, choosers were provided with basic information about both of the potential Dictators' years of education, age, and birth region. Prior to choosing "their" Dictator, we also elicited the subject's beliefs about how much each of the Dictators they were choosing between would allocate to them in the game (this is the source of the Dictator game beliefs data presented in Figure 1, Panel B). The subject's payoff in the game was the amount the selected Dictator actually transferred when he or she played the game.

The Choose-Your-Dictator game was played twice. In the first, "standard" version, choosers select a Dictator knowing that the potential Dictators do not have any information about them, i.e., the payoff will be from the standard Dictator game played by these Dictators. Through the lens of our model in section 2, this choice depends only on expected differences in average generosity b_i between the two potential Dictators, but not on their coethnic bias. In the second, "profiled" version, the choosers select a Dictator knowing that the Dictator will be provided with the choosers' own basic information profile (i.e., their years of education, age, and home region). The payoff now is from the profiled Dictator game.[5] In this second version of the game, where the Dictator has information about the likely ethnicity of the recipient, the chooser's selection depends on expected differences in $b_i q_i$ across the coethnic and non-coethnic Dictators. The difference in play across the "standard" and the "profiled" versions of the Choose-Your-Dictator game is of great conceptual interest for this study, as it helps to isolate the expected coethnic bias term q_i.

[5] In practice, it was impossible to perfectly match all characteristics from our set of pilot profiles to each and every subject in the actual lab sample, so the matching between the subject and the profile was done on coethnic home region alone. Given that we do not find that the years of education or age in the profiles are significantly correlated with the choices of subjects faced with these profiles, we do not believe that this "mismatch" of characteristics other than home region is consequential for the analysis.

To understand the extent of respondent preferences for a coethnic "Dictator", we implement an ordered logit analytical specification. The dependent variable is Y_{ip}, which takes on the following values for Dictator profiles p in the choice set of respondent i: $Y_{ip} = 0$ if profile p is not chosen by the subject, equals 1 if the subject is indifferent between profile p and the other profile in its choice set, and equals 2 if the subject chooses profile p. The "indifferent" option was included after we piloted the game and observed that large shares of participants claimed they were unable to choose between potential Dictators based on the basic profile information alone, an early hint about the low levels of coethnic bias in this population.

We specify the probabilities for observing each outcome value of Y_{ip} as a function of V_{ip}, which can be interpreted as respondent i's latent utility for profile p (Woolridge 2001). Using this framework, we carry out maximum likelihood estimation. The main specification estimates the difference between the valuation placed on a coethnic versus a non-coethnic profile across the "standard" and the "profiled" versions of the Choose-Your-Dictator game:

$$V_{ip} = \alpha_p + \beta_1 Coethnic_{ip} + \beta_2 Coethnic_{ip} * Profiled_{ip} + \beta_3 Coethnic_{ip} * Election_i + \beta_4 Coethnic_{ip} * Profiled_{ip} * Election_i + \varepsilon_{ip} \quad (8)$$

Here α_p is the set of profile fixed effects, $Coethnic_{ip}$ is an indicator variable for profile p being coethnic of the subject, and $Profiled_{ip}$ indicates whether the Dictator was playing the profiled version of the Dictator game, and thus had information about the subject's home region. In the Non-election round, β_1 estimates the degree of coethnic preference in the standard Choose-Your-Dictator game (equivalent to b_i generosity in our model above), and β_2 is the additional degree of coethnic preference in the profiled Choose-Your-Dictator game, which captures the degree of coethnic bias that the subject expects from the Dictator, which is equivalent to their expectation of the q_i coethnic preference

term in our model. Thus the estimate of β_2 is of central interest. The β_3 and β_4 terms capture any additional coethnic bias in the Election round. The error-term ε_{ip} has a Type I Extreme-value distribution. As a robustness check, we also condition on subject covariates, in this case interacted with the $Coethnic_{ip}$ term (since explanatory variables that do not vary across choice options cannot be estimated in a logit model).

3.6. Implicit Association Tests (IATs)

A potential concern with our approach is that game behavior may be affected by the desire of participants not to be seen to be discriminating along ethnic lines. We address this issue in two ways.

First, we collected data to ascertain whether participants were aware of our interest in ethnic discrimination. During piloting that we undertook before the start of the main data collection, we asked subjects – both at the midpoint of the lab sessions and at their conclusion – detailed questions regarding their understanding of the main focus of the experiment. The most common response was that the experiment was mainly about economic or business issues. Roughly equal numbers of subjects thought it was mainly about education or about "tribe" (i.e., ethnic issues). The proportion who believed that ethnic issues were a focus of the experiment rose slightly by the end of the lab, perhaps due to the nature of the Choose-Your-Dictator game, the final game, which asks participants to explicitly choose between two other players after being given information about those players' region of origin (among other characteristics), but still remained less than a quarter of all subjects. We are therefore confident that the vast majority of subjects did not understand the study's core research aims during the course of the labs; Appendix C contains further details.

In addition, we included an Implicit Association Test (IAT) as part of our main lab protocols. The objective of the IAT is to capture implicit biases of which respondents may not even be conscious. As such, the IAT results arguably allow us to avoid social

desirability bias and experimenter demand effects, both of which may affect game play if subjects are sufficiently aware of the study hypotheses and/or socialized against public displays of ethnic bias. This measure of implicit ethnic bias is complementary to our measures of explicit ethnic bias in the behavioral games.

The test is premised upon theories of learning and representation in psychology, and presumes that individuals find it easier to make the same behavioral response to concepts that are "strongly associated," i.e., about which they agree, than for concepts that are weakly associated (Nosek et al. 2007). In our lab, we focused on the degree to which respondents have a bias in favor of their own ethnic group.[6] Others have used related IAT methodologies to assess ethnic bias in Africa (Lowes et al 2015), gender bias in India (Beaman et al. 2009), and racial bias in the U.S. and elsewhere (Greenwald et al. 2003, Bertrand et al. 2005, Greenwald et al. 2009, Ogunnaike et al. 2010, Rooth 2010).

The IAT provides a measure of a respondent's automatic associations with certain Kenyan ethnic groups. Individuals were timed to assess how quickly they were able to associate coethnics with "good" traits and non-coethnics with "bad" traits. This is the "congruence" round of the IAT. In the "dissonance" round, coethnics are associated with bad traits and non-coethnics with good ones. A faster response time for congruence tasks compared to dissonance tasks implies that the subject has a more positive attitude toward coethnics (or a more negative attitude toward non-coethnics). Examples of several "screen shots" from the IAT are provided in Appendix C.

The outcome measure for the IAT is the within-respondent normalized difference in average reaction times (ART) between the dissonance (D) and congruence (C) tasks, called the d-score for subject i: $d_i = \frac{ART_i^D - ART_i^C}{\sigma_i}$, where σ_i is the standard deviation of a respondent's reaction times across all items. In keeping with the

[6] We also carried out a parallel "national" IAT to capture the strength of feelings toward Kenya, and plan to discuss those results in future studies (see Appendix E).

literature, we interpret a higher d-score as indicating a stronger coethnic bias in preferences in the ethnic IAT.

We first estimate the average level of ethnic bias in the IAT in the full sample. The correspondence between these patterns and the extent of coethnic bias in the games may shed light on the extent of experimenter demand effects in the games. We also estimate the effect of the priming treatments on the ethnic IAT d-score.

4. Results

4.1. Descriptive statistics

Descriptive statistics for the lab participants are presented in Table 1 (left column). Females are slightly more than half the sample (at 60%), respondents are 33 years old on average, and they have completed an average of 9.6 years of education, a schooling level that is typical for young Kenyan adults in urban areas but higher than national averages. In terms of employment, 29% of the sample report having a continuous source of wage income, 13% report self-employment, and 30% are unemployed (not shown). More than half of participants belong to the two largest ethnic groups, with 32% of the sample Kikuyu and 21% Luo, and the remainder from smaller groups.

Although not shown in the table, in Appendix B we present evidence that the randomization across the various priming interventions created treatment groups with similar observable characteristics along these dimensions.

Background data on subjects' political attitudes and activities in the survey data (collected at the end of the lab, Figure 2) corresponds closely to theories of ethnic-bloc voting in Africa (Horowitz 1985), and confirms the political salience of ethnicity in this sample. Approximately 84% of ethnic Luos report that they would support the coethnic frontrunner Raila Odinga in the then upcoming 2013 elections, against only 8% of ethnic Kikuyus, while 65% of Kikuyus state their support for Uhuru Kenyatta (versus just 1%

of Luos). Despite this polarization, the majority of respondents (63%) state that the 2007 post-election violence was not justifiable to any degree. Nearly 88% of individuals said they intend to vote in the 2013 election, similar to official national voter turnout.[7]

4.2. Estimating ethnic bias in preferences

As briefly mentioned in the introduction, there is no evidence for coethnic bias in either the Dictator game or the Public-good game. We elaborate on this finding in this section. Reproducing the result from Figure 1, the top row of Table 1 presents average transfers in both games in the full sample. For both games the differences in average transfers are small (35.6% to 35.4% in the Dictator game and 46.4% to 46.7% in the Public-good game, for Coethnic versus Non-coethnic giving, respectively) and not statistically significant.

The regression results confirm the finding. Given the large number of hypotheses we investigate, an important concern is the over-rejection of null hypotheses ("false positives") due to the problem of multiple inference (Anderson 2008). To mitigate this risk, we present two types of p-values in our analyses: first the regular or "per-comparison" p-value, which is appropriate when testing a single hypothesis, and second, the FWER-adjusted p-value, which captures the likelihood that at least one true null hypothesis is falsely rejected within a given set of hypotheses, using the free step-down resampling method described in Westfall and Young (1993).

The lack of a Non-coethnic arm in the Non-election round lab implies that the relevant test for coethnic bias in the Dictator game is the sum of the coefficient estimates on the Coethnic terms in both rounds minus the coefficient estimate on the Non-coethnic term in the Election round. The average coefficient on bias is just 0.6 percentage points and not statistically significant (Table 2, columns 1-2, p-value = 0.67). The large sample size implies that these are quite precisely estimated zeros: the 95%

[7] See the Kenya Independent Electoral and Boundaries Commission Election Results (accessed July 2015): http://www.iebc.or.ke/index.php/election-results

confidence interval on the overall Coethnic effect in the Election round ranges from –2.2 to +3.4 percentage points. For those in the priming control group (i.e., who received no priming to ethnic identity, political competition, or national identity), the effect is slightly larger in magnitude but not significantly different from zero (column 3, p-value = 0.48). In line with these findings, beliefs about the generosity of a coethnic Dictator are nearly identical to beliefs about non-coethnic Dictators and not significantly different from zero (Figure 1, Panel B, p-value = 0.51).

In the Public-good game, average coethnic bias in the Non-election round is approximately 1.57 percentage points but this difference is not statistically significant (Table 3, columns 1-2, p-value = 0.16). There is actually a significant reduction in coethnic bias in the Election round of the Public-good game, with an average effect of -3.67 percentage points (p-value = 0.027), which runs counter to the existing evidence that ethnic identity and preferences might become more salient closer to elections (Eifert et al. 2010, Michelitch 2015). Taken together, the overall degree of coethnic bias in the Election round is the sum of these two effects, and thus is small and negative (roughly -2 percentage points) and not statistically significant. In the priming control group, there are no significant coethnic bias effects in either lab round (column 3).

Subjects believe groups members will be equally generous in coethnic and ethnically mixed groups (Figure 1, Panel D), and results are unchanged when the dependent variable is "contributions minus beliefs," a measure of conscious free-riding on the part of the subject (Appendix B), as proposed in Fischbacher and Gächter (2010).

Even if overall coethnic bias is close to zero, a natural question is the degree of heterogeneity in the degree of this bias. This is illustrated in Table 1. The striking pattern that emerges is how little variation there is along the various dimensions that we pre-specified in the PAP. For instance, across both gender groups, the difference in average transfers in the Dictator game to coethnics and non-coethnics is less than one percentage point, and among subgroups defined by age, education, Ravens cognitive

score, and major ethnic groups, the differences are at most 1.2 percentage points. None of these differences is statistically significant at traditional confidence levels. A similar pattern holds for Public-good game contributions between coethnic and mixed groups, with coethnic bias estimates of at most 1.9 percentage points for any subgroup. Again, none of these differences are significant.

The evidence from the Dictator game and Public-good game are thus both consistent with little to no coethnic preference in this population as a whole, as well as for the major demographic subgroups. In terms of our model above, these patterns are consistent with values of q_i (coethnic bias in preferences) near zero for most subjects. Appendix E contains further descriptive statistics and histograms, as well as tests indicating that the null hypothesis of equality of the distributions of Coethnic and Non-coethnic transfers cannot be rejected in either game.

The Choose-Your-Dictator game also provides evidence on the degree of coethnic bias, although this game does so by shedding light on subjects' expectations about how much coethnic bias they will experience at the hands of others. Recall that in the standard Choose-Your-Dictator game, preference for a coethnic Dictator captures beliefs about the Dictator's likely generosity in general (the b_i term in the model), while the difference in preference for a coethnic Dictator in the standard versus the profiled version sheds light on how much coethnic bias the subject expects from others (their q_i).

The overall level of coethnic preference in the standard Choose-Your-Dictator game is minimal, with 27% choosing a coethnic versus 22% a non-coethnic, and half simply opting for "indifferent" in the Non-election round (Figure 3, left panel), and nearly identical patterns in the Election round (right panel). While seemingly small in magnitude, the effect is significant (Table 4, column 1) though not robust to the inclusion of covariates or focusing on the priming control group (columns 2-3).

The more important test from the point of view of understanding coethnic bias is the difference between these patterns and those that emerge in the profiled Choose-

Your-Dictator game, the version in which the Dictator is given information about the participant. In the profiled game, the proportion of subjects choosing a coethnic Dictator rises slightly, to 32%, as does the proportion who choose a non-coethnic (25%) in the Non-election round (Figure 3, left panel), and overall patterns of behavior in the profiled game are also presented in the top row of Table 1. As discussed above, the difference in behavior between the standard and profiled games is captured in the coefficient estimate on the Profiled Game * Coethnic term, and this effect is small and not statistically significant (Table 4, columns 1-3); taking into account the issue of multiple hypothesis testing, the FWER adjusted p-value is close to one (at 0.993). Nor does this effect emerge in the Election round (columns 1-3).

In other words, there is no evidence that subjects are more likely to choose a coethnic Dictator when the Dictator has information about their home region (and thus a way to make an inference about their ethnic background), relative to when the Dictator does not have this information. This is consistent with the previous findings since there is no evidence for coethnic bias in behavior in the Dictator game (Table 2). Expectations and actions regarding coethnic bias are thus aligned among our subjects.

4.3. Priming impacts

None of our priming treatments had a statistically significant effect on behavior in either the standard Dictator game (Table 5, column 1) or the standard Public-good game (column 4), and these effects did not differ significantly in the Election round (columns 2 and 5, respectively).

As noted, the ethnic-political prime was only included in the Election round. This is because, after finding no effect of the ethnic-cultural prime in the 2012 Non-election round, we hypothesized that this might be because our priming approach was too subtle. Hence, we decided to introduce a more "blatant" ethnic priming treatment in the Election round. In contrast to the ethnic-cultural prime, which sought to enhance

the salience of ethnic identity by emphasizing cultural aspects of ethnic identity, this new treatment directly and overtly primed subjects to the link between ethnic identity and political outcomes. For example, one of the questions asked: *"Which of the following ethnic groups controls the largest share of cabinet positions?"* Yet this prime also has no significant effect on transfers in either game (Table 5, columns 2 and 5).

There is also little evidence of priming effects in the profiled Dictator game, Public-good game or Choose-Your-Dictator game; nor does the priming effect differ across the Non-election and Election rounds (detailed results are in Appendix B, Tables S1, S2, and S3; the full set of prespecified results are in Appendix E). Few of the relevant differential priming effect estimates are statistically significant in the per-comparison sense, and almost none survive the multiple testing adjustment.

The bottom line is that the priming interventions, which aimed to raise the salience of ethnicity, political competition and national identity did little to change behavior in the lab activities, including the degree of coethnic bias. We interpret this as broadly consistent with the above findings regarding both the lack of a coethnic bias effect in our sample and the absence of a difference in the extent of bias across the two lab rounds. The results suggest that coethnic bias does not appear to lie "just below the surface" in a way that might be triggered by priming subjects to ethnicity or political competition. They also suggest that priming subjects to their membership in a superordinate national identity does not affect their (already low) levels of ethnic bias.

4.4. Implicit association test (IAT) results

A final measure of coethnic bias can be gleaned from subjects' behavior in the IAT. We find a statistically significant and positive coethnic bias effect in this test, with a magnitude of 0.079 standard deviation units (p-value < 0.01, Table 6, column 1). This degree of bias is considered "small" in the related IAT research literature in psychology (Cohen 1988). For instance, studies of racial bias in the United States – i.e., of the bias

whites holds against blacks – find estimates that are roughly six times as large, in the range from 0.45 to 0.52 (Nosek et al. 2007), while the average gender bias regarding political leadership in India is roughly twice as large, at 0.11 to 0.15 (Beaman et al. 2009). Lowes et al (2015) report d-score values of 0.14 in a study of ethnic bias in rural Congo, nearly twice as large as our estimates.

There is some evidence that the political competition priming treatment increased ethnic bias in the IAT (Table 6, columns 3-4), which might be interpreted as consistent with the findings in Eifert et al. (2010). However, this effect is only marginally significant once multiple testing considerations are accounted for (column 5). None of the other priming interventions affected implicit ethnic bias.[8]

4.5. Election proximity effects

While not the main focus of the study, we document an interesting set of patterns regarding the effects of lab proximity to the 2013 Kenya elections. These tests were pre-specified in the pre-analysis plans, so below we present p-values both in the typical per-comparison sense, and in terms of FWER multiple testing adjustments.

Average transfers in the Dictator game are markedly lower in the Election round compared to the Non-election round, with a difference of over 5 percentage points (Appendix Figure S2, left panel). This difference is statistically significant (per-comparison p-value < 0.01, Table 2 columns 1-3), and also survives the multiple testing adjustment (FWER p-value = 0.008, column 4). This is consistent with an increase in egoism (b_i) in the theoretical framework above.

Patterns are similar, although somewhat more muted, in the Public-good game, with an average effect of roughly three percentage points (Appendix Figure S2, right

[8] Note that since the IATs were administered only in the Election round we are unable to assess whether implicit ethnic bias is different in proximity to an election.

panel). This effect is significant at 90% confidence in the traditional sense (Table 3, columns 1-3) but is not significant after FWER adjustment (p-value = 0.321, column 4).

Other scholars have documented that real-world economic conditions matter for lab generosity (e.g., Fisman et al. 2014), and some scholars have argued specifically that competitive environments might lead to a breakdown in adherence with cooperation norms (Falk and Szech 2013, Michelitch 2015). However, we are unaware of other research that documents systematically lower levels of generosity in election periods compared to non-election periods. Further research would be useful to understand how much this pattern generalizes across settings and elections.

Interpretation of this finding is less straightforward than many other results in this paper since the Non-election round versus Election round comparison is not based on experimental variation in the lab. At its base, this is a "time effect" between the two lab rounds, and any other changes to the Kenyan setting, the study sample, or the Busara lab space itself would contribute to effects, not just proximity to the election alone. At best, then, this finding should be seen as provocative but not definitive.

With that in mind, there are several reasons to think that these round effects might in fact be capturing an election proximity effect, at least in part. First, there were no major economic "shocks" during the study period in terms of prices (see Appendix Figure S1) or asset values, making it seem unlikely that a macroeconomic shock like that studied in Fisman et al. (2014) is driving the result. In contrast, the political climate changed dramatically between rounds: the survey we administered indicates that in the Non-election round, just 25% (14%) of our subjects reported having recently attended a campaign rally (received cash from a candidate), but this more than doubles to 61% (31%) by the time of the January to February 2013 Election round.

The Busara Center also used comparable subject recruitment methods during our 2012 Non-election round and our 2013 Election round (and none of the subjects in the Election round had participated in the earlier labs), which suggests that data is

comparable across rounds. Despite common recruitment practices, however, there are some differences in the observed characteristics of lab subjects across the two rounds. While the samples for the Non-election round and Election round are statistically indistinguishable in terms of average age and education, they do differ significantly in the share of females (which is higher in the Election round) and the average Ravens score (lower in the Election round). The Election round sample also has a somewhat smaller share of Kikuyus (see Appendix B, Table S4). Because of this partial imbalance along observables across the lab rounds, we also estimated the main regressions specifications including these individual demographic characteristics as covariates (Table 2, column 2 and Table 3, column 2), and the estimated Election round effect actually becomes somewhat larger. Yet it remains impossible for us to rule out that selection of subjects along some unobservable dimension – for instance, their degree of political activism, which could have affected their willingness to participate in a lab activity rather than work on a political campaign – did differ across lab rounds, and this calls for further caution in the interpretation of the Election round results.

4.6. Assessing the presentation of results using the pre-analysis plan

Given the space constraints of a journal article, it was not possible to present every result that was pre-specified in our pre-analysis plans. This raises the possibility that we might have emphasized the subset of results that were more (or less) statistically significant, or more consistent with our theoretical priors. One means – to our knowledge, novel – of assessing whether we might have done this is to plot the distribution of multiple testing FWER adjusted p-values presented in the paper's main tables (N=86 tests) together with the distribution of FWER adjusted p-values for all hypothesis tests specified in the pre-analysis plans (N=763 in total).

We do this in Figure 4. The distributions are broadly similar, but the main study tables do appear to slightly over-report statistically significant estimates (those with p-

values less than 0.05) for both the unadjusted p-values (Panel A) and the FWER adjusted p-values (Panel B), and somewhat under-report those with p-values close to one. Consistent with this visual inspection, the Kolmogorov-Smirnov test on the equality of these two distributions is rejected at 95% confidence for both the unadjusted and the FWER adjusted p-values.

Of course, since the central implication of our empirical analysis is a null result, namely, the lack of evidence for coethnic bias in preferences across multiple lab experiments and measures, the slightly disproportionate emphasis on statistically significant results apparent in Figure 4 does not appear to be driving our main conclusions. Rather we think it likely reflects our reporting of the Election round results discussed above, which feature in Tables 2, 3, and 5, and our decision to relegate a large number of null priming treatment estimates and interaction effects to the appendix.

5. Conclusion

It is well-documented that ethnic differences are associated with negative outcomes in many settings, but the underlying mechanisms remain poorly understood. Using a large sample of over 1,300 participants and a rich set of experimental methods and measures in a Kenyan setting in which ethnic divisions are thought to be particularly important, we find little evidence of ethnically biased behavior or preferences.

These findings challenge the conventional wisdom about the centrality of ethnic preferences in explanations for the negative association between ethnic diversity and economic and political outcomes in Africa. They suggest that other factors—such as technologies that facilitate cooperation or the enforcement of norms within ethnic groups—may be even more important than has previously been thought. A key implication is that efforts to dampen ethnic divisions by changing ethnic attitudes may not be enough. Rather, institutional and policy reforms that facilitate the flow of

information across groups and limit the ability of elites to mobilize the population along ethnic lines may prove more effective in ameliorating politicized social divisions.

Our findings are in apparent tension with the results of Hjort's (2014) important recent study of ethnic bias that also uses Kenyan data. Hjort takes advantage of the random assignment of workers to teams on a flower farm to study whether within-team productivity is lower when those teams are ethnically diverse. He finds that it is, and argues that this is due to discrimination by team members on behalf of coethnics. Hjort shows that this diversity effect is magnified after the 2007-08 election violence.

There are several possible ways to account for the differences between these findings and our own. The most immediate is that, though attributed to ethnic preferences, the negative diversity effects in Hjort may in fact be caused in part by other mechanisms. For example, the fact that Hjort shows that modifications in contractual details—namely, moving to group-based pay on work teams—mitigates much of the negative effect suggests that institutional factors may be critical. Hjort's design also makes it impossible to rule out the possibility that shared ethnicity could provide a technology that facilitates team production. One of the strengths of our laboratory approach is that it allows us to focus exclusively on the preferences mechanism and to rule out alternative channels.

It is also important to note that while political tensions around the 2007 elections (and in their aftermath) were extremely high, the 2013 national elections were conducted in a much more peaceful atmosphere. This was in part due to measures taken by the Kenyan government and political candidates, as well as to the efforts of non-governmental actors to emphasize the importance of avoiding intergroup violence. Hence the expectation that our period of study would be one of particularly high ethnic tensions – and, by implication, that we would find high levels of ethnic bias in the lab round proximate to the elections – may have been somewhat misplaced.

Another possible source of the discordant findings with Hjort's (2014) study lies in differences in our samples. Both Hjort's sample and ours feature many of the same ethnic groups and wide range of young and middle-aged adults, so those characteristics are unlikely to explain differences. But it remains possible that ethnic bias is simply stronger in his rural agricultural setting than in our urban sample.

Given the strictly urban nature of our subject pool, our study is not particularly well suited to fully assessing this possibility. However, we are able to make some progress by comparing the behavior in our sample of subjects who were long-time Nairobi residents with the behavior of those who had moved to the city relatively recently. While only suggestive, this is a meaningful test to the extent that individuals gradually assimilate to local norms over time, such that those who have moved to Nairobi relatively recently have preferences more similar to those who remain in rural areas. Such an interpretation is consistent with research that finds that migrants gradually adapt their social preferences and behaviors over time to the norms in their new location (Henrich et al. 2006, Jang and Lynham 2015, Laitin 1998).

This analysis was not specified in our pre-analysis plan, and is thus more speculative than our other empirical results. While we would ideally focus on those individuals who had just moved to the city, there are relatively few such individuals in our lab sample: just 2 percent had lived in Nairobi for less than one year. However, a sizeable proportion had moved to Nairobi within 5 years (roughly 20% of subjects) and even more in the last 10 years (40%). We examine the behavior of this latter group in comparison with that of longer-term urban residents. Consistent with the hypothesis that there might be an important urban-rural divide when it comes to ethnic preferences, we find statistically significant differences in behavior between long-term Nairobi residents and more recent arrivals. While individuals who have resided in Nairobi for more than a decade exhibit no evidence of coethnic bias in their behavior, more recent migrants are significantly more generous toward their coethnics in the

Election round for the profiled Dictator game, giving an average of 4.2 percentage points (standard error 1.8) more to coethnics. Coethnic bias in the Dictator game is even larger among those who had resided in Nairobi for at most 5 years, with an average increase of 6.2 percentage points (standard error 2.4).

Thus migrants from rural settings appear to gradually assimilate to a new set of norms regarding interethnic cooperation and ethnic preferences over time, perhaps in part due to the far more extensive interactions they have with non-coethnics in Kenya's cosmopolitan capital. Further evidence comes from our political attitudes survey, which shows rising ethnic identification (relative to other dimensions of individual identity) in the 2013 Election round among those who have lived in Nairobi for less than 10 years (coefficient estimate 8 percentage points, significant at 90% confidence) but not among longer-term Nairobi residents.

Obviously, these findings cannot be taken as causal for many reasons, including the fact that we lack rural lab participants. We also cannot rule out that the differences mentioned above are driven by selective urban migration patterns over time rather than the true effect of the length of urban residency, and furthermore, as noted, these analyses were not pre-specified. Nevertheless, the suggestion that ethnic preferences may be weaker in urban than in rural settings is provocative—in part because it would call for a reevaluation of the well-known variant of modernization theory that posits a positive relationship between urbanization and ethnic divisions (e.g., Bates 1983).

The possibility that ethnic preferences in contemporary Africa might indeed be weaker in urban areas does find at least some support in the empirical literature. In Kampala, Uganda (an urban setting similar to the one we study), Habyarimana et al (2007) also find no evidence of ethnic bias in the Dictator and Public-good games. Lowes et al (2015), meanwhile, do find evidence of ethnic bias in the IAT in a rural sample in Congo. In a cross-national Afrobarometer sample, Robinson (2014) finds that urban dwellers are more likely than their rural counterparts to privilege their national

identity over their ethnic identity—a result one step removed from, but consistent with, weaker ethnic preferences in urban areas.

More is at stake here than simply helping to reconcile seemingly disparate findings in our study versus Hjort (2014) and other recent literature. Our suggestive findings also provide new insight into the future trajectory of ethnic preferences and identification in Sub-Saharan Africa, given the rapid urbanization taking place in many countries and the growing cross-ethnic interactions that accompany it. Other scholars have pointed to the role that expanding democratization may play in dampening ethnic favoritism (Burgess et al. 2015). Gaining a better understanding of the role played by rapid urbanization, as well as by the major institutional reforms accompanying democracy, will be an important goal for future scholarship on African ethnic divisions.

References

Adida, Claire L, David D. Laitin and Marie-Anne Valfort. 2015. Religious homophily in a secular country; Evidence from a voting game in France. *Economic Inquiry* 53(2).

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114 (4): 1243–84.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43 (3): 762–800.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Barkan, Joel D., and Michael Chege. 2008. "Decentralising the State: District Focus and the Politics of Reallocation in Kenya." *Journal of Modern African Studies* 27 (03): 431.

Bates, Robert. 1983. "Modernization, Ethnic Competition and the Rationality of Politics in Contemporary Africa." In Donald Rothchild and Victor Olorunsola, eds. *State Versus Ethnic Claims: African Policy Dilemmas*. Westview Press: 152-171.

Beaman, L, R Chattopadhyay, E Duflo, R Pande, P Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of Economics* 124 (4): 1497–1540.

Bertrand, M, D Chugh, S Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95 (2): 94–98.

Bowles, Samuel and Herb Gintis. 2004. "Persistent Parochialism: Trust and Exclusion in Ethnic Networks." *Journal of Economic Behavior and Organization* 55: 1-23.

Burgess, R, R Jedwab, E Miguel, A Morjaria, G Padró i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review* 105 (6): 1817–51. doi:10.1257/aer.20131031.

Cappelen, AW, AD Hole, EØ Sørensen, B Tungodden. 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review* 97 (3): 818–27.

Cappelen, AW, J Konow, EØ Sørensen, B Tungodden. 2013. "Just Luck : An Experimental Study of Risk Taking and Fairness." *American Economic Review* 103 (4): 1398–1413. doi:10.1257/aer.103.4.1398.

Carlson, Elizabeth. 2015. "Ethnic Voting and Accountability in Africa: A Choice Experiment in Uganda." *World Politics* 67 (02): 353–85.

Casey, K, R Glennerster, and E Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics* 127 (4): 1755–1812.

Cederman, Lars-Erik, and Luc Girardin. 2007. "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies." *American Political Science Review* 101: 173.

Cederman, LE, NB Weidmann, KS Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Political Science Review* 105 (03): 478–95. doi:10.1017/S0003055411000207.

Coffman, LC, M Niederle. 2015. "Pre-Analysis Plans Have Limited Upside, Especially Where Replications Are Feasible." *Journal of Economic Perspectives* 29 (3): 81–98.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. *Statistical Power Analysis for the Behavioral Sciences*. Vol. 2nd. doi:10.1234/12345678.

Dionne, K. Y. 2014. "Social Networks, Ethnic Diversity, and Cooperative Behavior in Rural Malawi." *Journal of Theoretical Politics*, November.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112 (4): 1203–50.

Eifert, Benn, Edward Miguel, and Daniel N. Posner. 2010. "Political Competition and Ethnic Identification in Africa." *American Journal of Political Science* 54 (2): 494–510.

Falk, Armin, and Nora Szech. 2013. "Morals and Markets." *Science* 340 (6133): 707–11.

Fischbacher, U, S Gächter. 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *American Economic Review* 100 (1): 541–56.

Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2014. "How Did Distributional Preferences Change During the Great Recession?" *NBER Working Paper 20146*.

Gaertner, Samuel L. and John F. Dovidio. 2000. *Reducing Intergroup Bias: The Common Ingroup Identity Model*. Psychology Press.

Greenwald, AG, BA Nosek, MR Banaji. 2003. "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm." *Journal of Personality and Social Psychology*. doi:10.1037/0022-3514.85.2.197.

Greenwald, AG, CT Smith, N Sriram, Y Bar-Anan, BA Nosek. 2009. "Implicit Race Attitudes Predicted Vote in the 2008 U.S. Presidential Election." *Analyses of Social Issues and Public Policy* 9 (1): 241–53. doi:10.1111/j.1530-2415.2009.01195.x.

Grossman, Shelby and Dan Honig. 2015. "These Are My People: Ethnic and Class Based Discrimination in Lagos." *Working paper*.

Hardin, Russell. 1995. *One For All: The Logic of Group Conflict*. Princeton Univ. Press.

Habyarimana, J, M Humphreys, DN Posner, JM Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101 (4): 709–25. doi:10.1017/S0003055407070499.

Henrich, J, R McElreath, A Barr, J Ensminger, C Barrett, A Bolyanatz, JC Cardenas, et al. 2006. "Costly Punishment across Human Societies." *Science* 312 (5781): 1767–70.

Hjort, J. 2014. "Ethnic Divisions and Production in Firms." *Quarterly Journal of Economics* 129 (4): 1899–1946. doi:10.1093/qje/qju028.

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

Jang, Chaning, John Lynham. 2015. "Where Do Social Preferences Come From?" *Working Paper*.

Jeon, Sangick. 2013. "Managing Ethnic Diversity: Experimental Evidence from Kenya." *Working Paper, Stanford University*.

Kramer, RM, MB Brewer. 1984. "Effects of Group Identity on Resource Use in a Simulated Commons Dilemma." *J. of Personality and Social Psychology* 46(5): 1044-57.

Laitin, David D. 1998. *Identity in Formation: The Russian-Speaking Populations in the Near Abroad*. Ithaca, NY: Cornell University Press.

Lieberman, Evan S. and Gwenyth McClendon. 2013. "The Ethnicity-Policy Preference Link in Sub-Saharan Africa." *Comparative Political Studies* 46 (5): 574-602.

Loewen, PJ, G Eady, J Fowler. 2014. "Social Distance Is More Important than Coethnicity for Cooperative Behavior: Results from Experiments in Rural Malawi." *Working Paper*.

Lowes, S, N Nunn, JA Robinson, J Weigel. 2015. "Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)." *American Economic Review: Papers & Proceedings* 105 (5): 340-345: doi: 10.1257/aer.p20151075.

Marx, B, TM Stoker, T Suri. 2015. "There Is No Free House: Ethnic Patronage in a Kenyan Slum." *Working paper.*

Melson, R, H Wolpe. 1970. "Modernization and the Politics of Communalism: A Theoretical Perspective." *American Political Science Review* 64(December): 1112-1130.

Michelitch, Kristin. 2015. "Does Electoral Competition Exacerbate Interethinc or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining." *American Political Science Review* 109 (1): 43-61.

Miguel, E., C. Camerer, K. Casey, et al. 2014. "Promoting Transparency in Social Science Research." *Science* 343 (6166): 30–31.

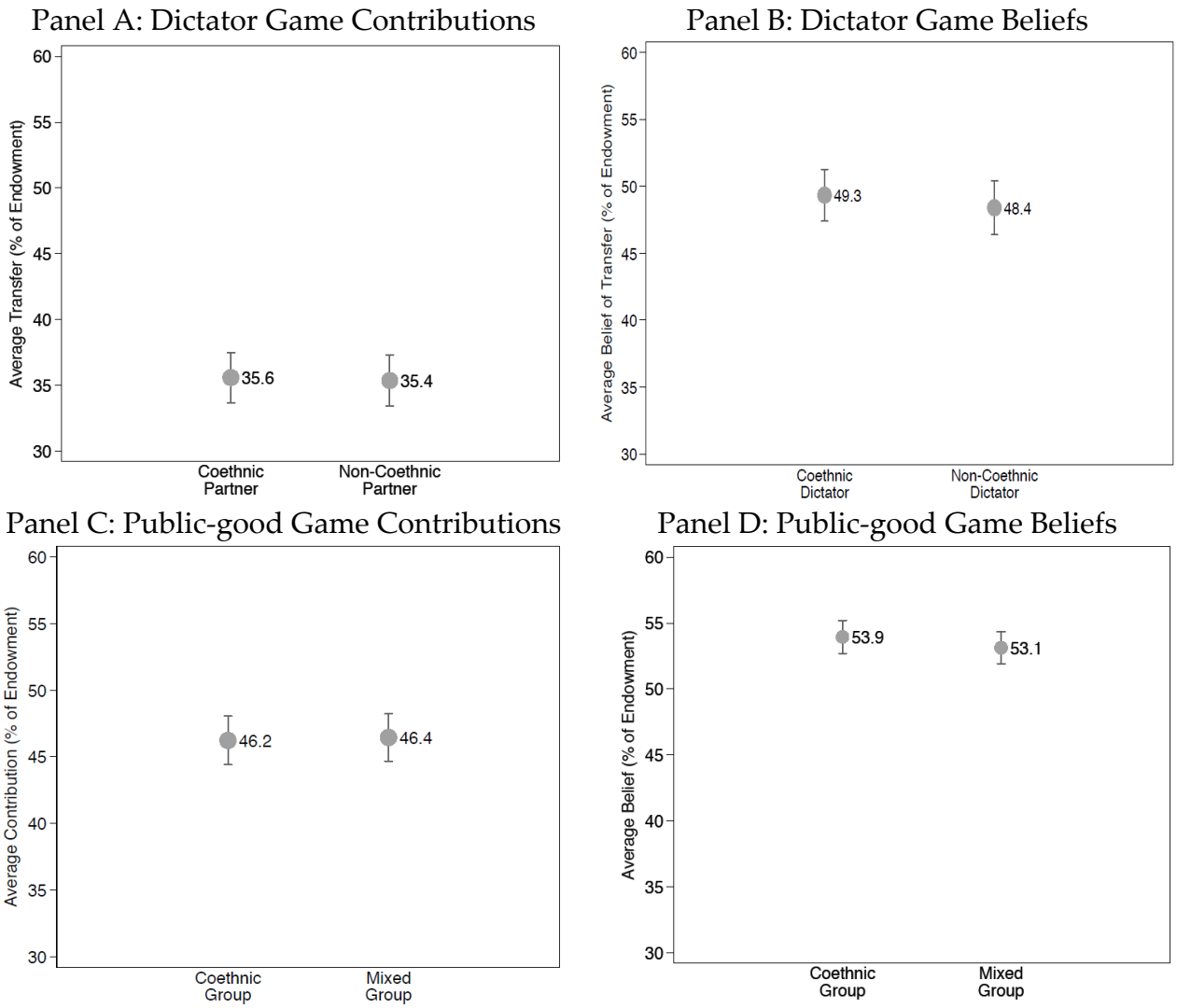
Miguel, E, MK Gugerty. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics* 89 (11-12): 2325–68.

Nosek, BA, FL Smyth, JJ Hansen, et al. 2007. "Pervasiveness and Correlates of Implicit Attitudes and Stereotypes." *European Review of Social Psychology* 18 (1): 36–88.

Ogunnaike, Oludamini, Yarrow Dunham, and Mahzarin R. Banaji. 2010. "The Language of Implicit Preferences." *Journal of Experimental Social Psychology* 46 (6): 999–1003.
Raven, John. 2008. "The Raven Progressive Matrices Tests: Their Theoretical Basis and Measurement Model." In *Uses and Abuses of Intelligence. Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*, edited by John Raven and Jean Raven, 17–68. Unionville, NY: Royal Fireworks Press.
Robinson, A. 2014. "National Versus Ethnic Identification in Africa: Modernization, Colonial Legacy and the Origins of Territorial Nationalism." *World Politics* 66 (4).
Rooth, Dan-Olof. 2010. "Automatic Associations and Discrimination in Hiring: Real World Evidence." *Labour Economics* 17 (3): 523–34. doi:10.1016/j.labeco.2009.04.005.
Shayo, M. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class and Redistribution." *American Political Science Review* 103 (2): 147-174.
Snyder, J. 2000. *From Voting to Violence: Democratization and Nationalist Conflict*. New York: W.W. Norton.
Vigdor, JL. 2002. "Interpreting Ethnic Fragmentation Effects." *Economics Letters* 75 (2).
Voors, MJ, EEM Nillesen, P Verwimp, et al. 2012. "Violent Conflict and Behavior: A Field Experiment in Burundi." *American Economic Review* 102 (2): 941–64.
Westfall, Peter H., and S. Stanley Young. 1993. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. New York: John Wiley & Sons, Inc.
Woolridge, J. M. 2001. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

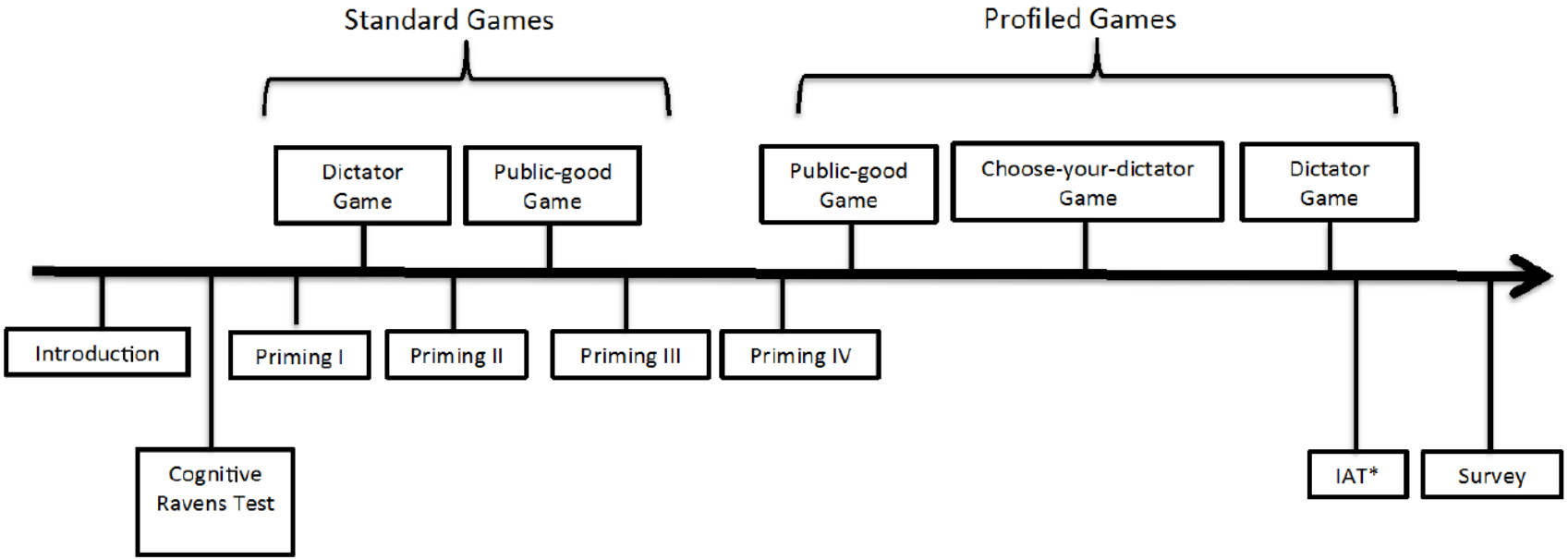
Tables and Figures:

Figure 1: Coethnic Bias in the Dictator Game and Public-good Game



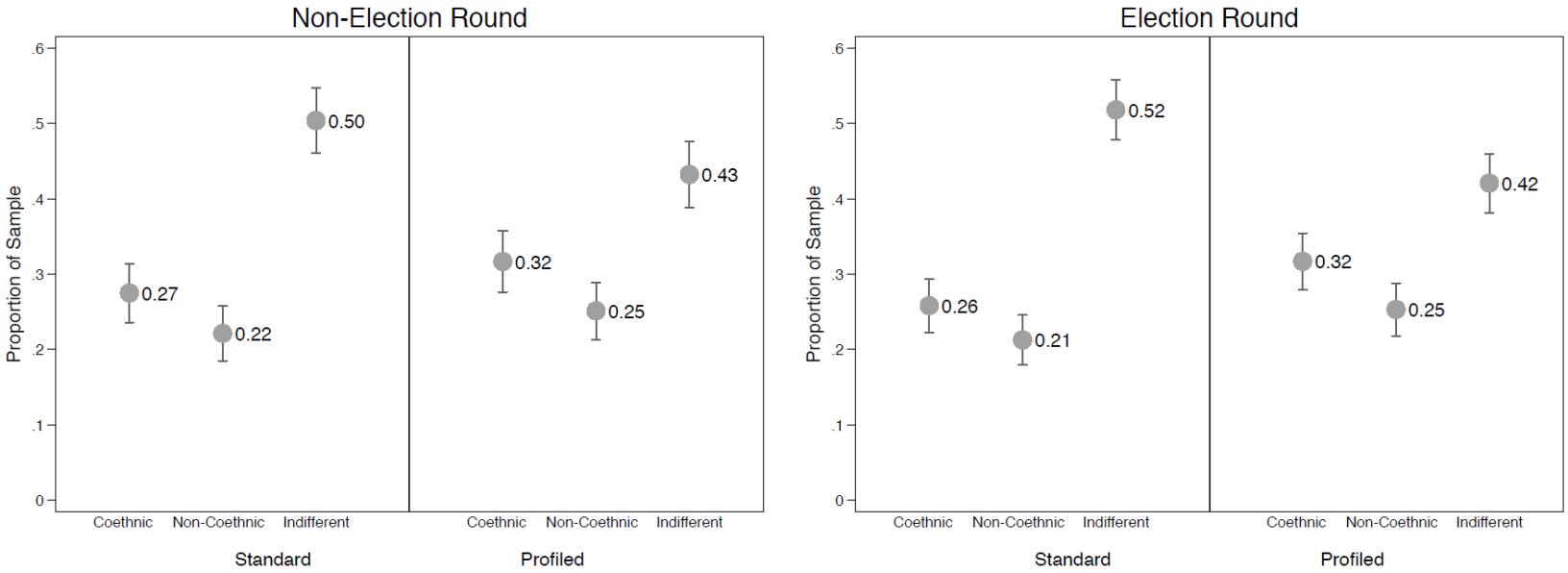
Notes: Sample averages and 95% confidence intervals for Dictator game contributions in profiled games for Coethnic vs. Non-Coethnic Transfers (Panel A), beliefs about Dictator game contributions from the profiled Choose-Your-Dictator game (Panel B), for Public-good game contributions in profiled games for Coethnic vs. Mixed Groups (Panel C), and for Public-good game beliefs about others' contributions (Panel D). The Dictator game data in Panels A and B is from the Election round (January-February 2013), the only time the complete profiled game data was collected. The Public-good game data is pooled from both the Non-Election round (July/August 2012) and the Election round, since the complete profiled game data was collected in both. The p-value of Coethnic = Non-Coethnic for Panel A is 0.87. The p-value of Coethnic = Non-Coethnic for Panel B is 0.51. The p-value of Coethnic = Mixed for Panel C is 0.86. The p-value of Coethnic = Mixed for Panel D is 0.36. In the Dictator game, participants had an endowment of 50 Kenya Shillings, and in the Public-good game participants were endowed with 60 Kenya Shillings.

Figure 2: Experimental Laboratory Game Structure



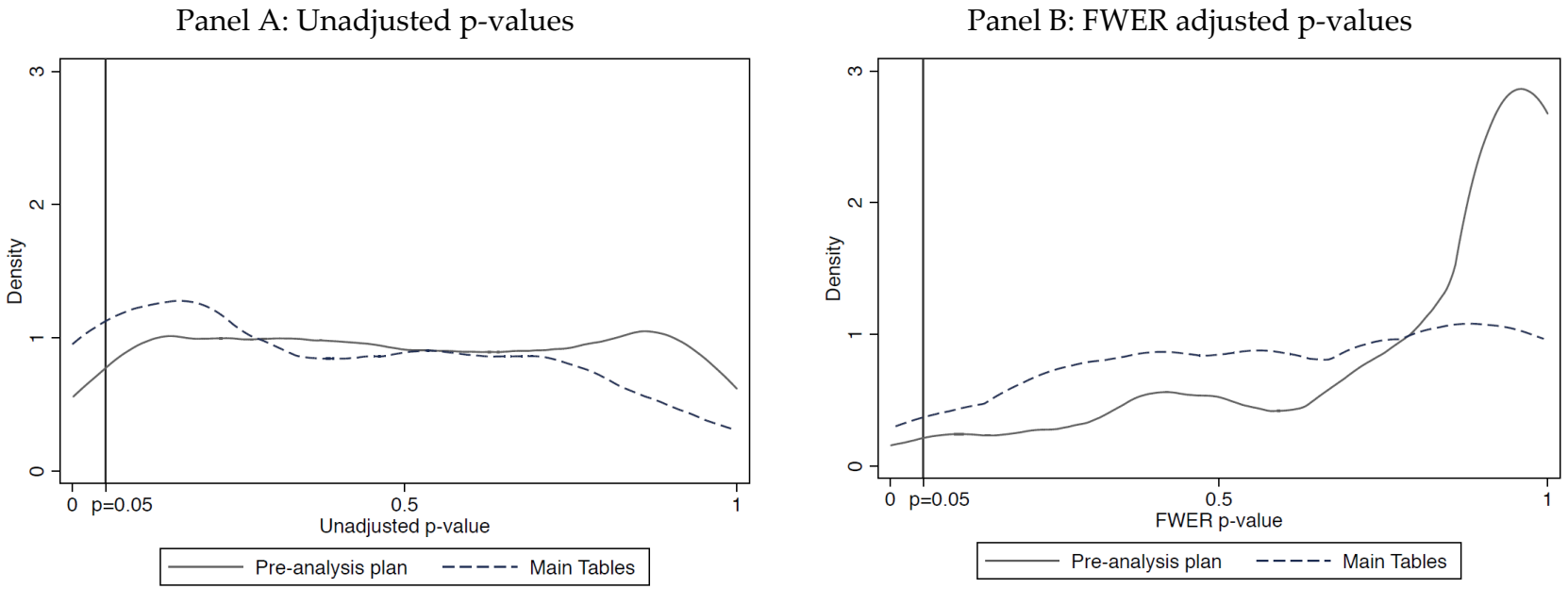
Notes: The Non-Election Round took place during July and August 2012 with a sample size of N=608. The 2013 Election Round took place during January and February 2013 with a sample size of N=754. The average lab time was roughly 3 hours. *Indicates that we only have this data for the Election Period.

Figure 3: Coethnic Bias in the Choose-Your-Dictator Game



Notes: Sample averages and 95% confidence intervals for Standard and Profiled Choose-Your-Dictator (CYD) games during the Non-Election round (left Panel) and Election round (right Panel). In the standard CYD game, the Dictator had no information about the chooser, as in a standard Dictator game. In the profiled CYD game, the Dictator had some information about the chooser, as in the profiled Dictator game in the lab. The chooser had profile information about the Dictator for both versions of the CYD game. The Kenya Non-Election round took place in July/August 2012, and the Kenya Election round took place in January/February 2013. Choosers could opt for a coethnic Dictator, a non-coethnic Dictator, or could state their indifference between the two, in which case they were randomly assigned one of the two Dictators.

Figure 4: Distribution of P-values in Pre-Analysis Plans versus Main Tables



Notes: The distribution of p-values from the pre-analysis plan includes p-values for all hypothesis tests discussed in both pre-analysis plans (Non-election round, Election round). The distribution of p-values from the main tables includes all tables (Tables 1 through 6) in this article. The dark vertical line denotes statistical significance at the standard 95% confidence level. Panel A presents p-values unadjusted for multiple testing. Panel B presents FWER adjusted p-values. Kolmogorov-Smirnov tests on the equality of the two distributions yields a p-value of 0.014 in Panel A, and 0.005 in Panel B.

Table 1: Descriptive Statistics on the Sample and Average Game Play

	Dictator Game Transfer (Percent of Endowment)		Public-Good Game Contribution (Percent of Endowment)		Choose-Your-Dictator Game (Dictator Choice)	
	Coethnic	Non-Coethnic	Coethnic Group	Mixed Group	Coethnic	Non-Coethnic
Demographics						
Full Sample	35.6	35.4	46.2	46.4	31.7	25.2
Female (60% of sample)	36.8	35.8	44.9	46.2	27.3	24.3
Male (40% of sample)	33.4	34.6	48.0	46.8	38.0	26.4
Age (mean = 33.0 years)						
Below Median	33.3	32.1	41.2	43.4	32.9	27.2
Median or Above	37.8	38.5	48.0	47.5	30.5	23.2
Education (mean = 9.6 years)						
Below Median	36.5	35.4	46.9	46.2	27.6	26.4
Median or Above	34.4	35.4	45.6	46.6	35.0	24.3
Ravens Score (mean = 0.0)						
Below Median	37.5	37.5	47.3	48.7	28.5	25.2
Median or Above	33.2	32.8	45.2	44.3	34.7	25.2
Ethnic Affiliation						
Kikuyu (32% of sample)	36.3	36.1	45.5	47.4	26.7	25.7
Luo (21% of sample)	33.4	34.5	42.7	43.6	37.2	23.8
Observations	593	593	967	981	981	

Notes: Pooled data from the Non-Election Round (July-August 2012) and Election Round (January-February 2013). Values are averages, presented in percent terms. For the Dictator game, only the Election Round averages are displayed, as there is incomplete information for the profiled version of the Dictator game in the Non-election Round. Respondents with Kamba ethnicity are excluded from analysis, as pre-specified in the pre-analysis plan.

Table 2: Dictator Game Transfers, in Standard and Profiled Games

	Full Sample		No Prime	FWER p-value
	(1)	(2)	(3)	(4)
Coethnic Recipient	-1.51	-1.44	-1.85	0.518
	(1.22)	(1.22)	(2.27)	
Election Round	-5.21***	-5.98***	-6.52**	0.008
	(1.41)	(1.43)	(2.60)	
Election Round * Coethnic Recipient	0.70	0.63	-0.83	0.786
	(1.88)	(1.87)	(3.46)	
Election Round * Non-coethnic Recipient	-1.41	-1.42	-4.54*	0.153
	(1.42)	(1.42)	(2.61)	
Covariates	No	Yes	No	
Observations	2881	2881	748	
Recipient: Coethnic + Election Round * (Coethnic - Non-coethnic)	0.60	0.60	1.86	
[P-value]	[0.67]	[0.67]	[0.48]	

Notes: The dependent variable is the transfer in the Dictator game (in percent of the endowment). Pooled data from the Non-election Round (July-August 2012) and Election Round (January-February 2013). Standard errors in parentheses are clustered at the individual level. P-values: * p<0.10, ** p<0.05, *** p<0.01. Covariates include ethnicity indicators, a gender indicator, education controls, and the Raven's test score. FWER p-values are simulated as described in the pre-analysis plan, for column (3). There was no Non-coethnic profile in the Dictator game during the Non-election Round, hence the absence of a direct "Non-Coethnic Recipient" term. The F-test in the bottom row tests the hypothesis that the average level of coethnic bias in the Election round was zero. Each individual was endowed with 50 Ksh (approx. 0.60 USD). All specifications exclude ethnic Kamba subjects, as specified in the pre-analysis plan.

Table 3: Public-good Game Contributions, in Standard and Profiled Games

	Full Sample		No Prime	FWER p-value
	(1)	(2)	(3)	(4)
Coethnic Group	2.22*	2.22*	-1.73	0.984
	(1.33)	(1.33)	(2.54)	
Mixed Group	0.65	0.65	0.32	0.988
	(1.18)	(1.18)	(2.29)	
Election Round	-2.97*	-3.95**	-6.03*	0.321
	(1.76)	(1.81)	(3.37)	
Election Round * Coethnic Group	-1.67	-1.68	2.00	0.984
	(1.89)	(1.89)	(3.39)	
Election Round * Mixed Group	2.00	2.00	2.30	0.984
	(1.85)	(1.85)	(3.59)	
Covariates	No	Yes	No	
Observations	2939	2939	763	
Coethnic Group - Mixed Group	1.57	1.57	-2.05	
[P-value]	[0.16]	[0.16]	[0.35]	
Election Round * (Coethnic Group- Mixed Group)	-3.67	-3.67	-0.30	
[P-value]	[0.027]	[0.027]	[0.93]	

Notes: The dependent variable is the contribution in the Public-good game (in percent of the endowment). Pooled data from the Non-election Round (July-August 2012) and Election Round (January-February 2013). Standard errors in parentheses are clustered at the individual level. P-values: * p<0.10, ** p<0.05, *** p<0.01. Covariates include ethnicity indicators, a gender indicator, education controls, and the Raven's test score. FWER p-values are simulated as described in the pre-analysis plan, for column (3). Each individual was endowed with 60 Ksh (approx. 0.70 USD). The first F-test tests the hypothesis that the average level of coethnic bias across both the Non-election round and the Election round is zero, and the second F-test tests the hypothesis that the difference in coethnic bias across the Non-election round and the Election round is zero. All specifications exclude ethnic Kamba subjects, as specified in the pre-analysis plan.

Table 4: Choose-Your-Dictator Game Choices, in Standard and Profiled Games

	Full Sample		No Prime	FWER p-value
	(1)	(2)	(3)	(4)
Coethnic Profile	0.25**	0.19	0.071	0.979
	(0.11)	(0.15)	(0.23)	
Profiled Game * Coethnic	0.048	0.048	0.021	0.993
	(0.11)	(0.11)	(0.21)	
Election Round * Coethnic	-0.059	-0.0013	0.15	0.979
	(0.11)	(0.12)	(0.23)	
Election Round * Profiled Game * Coethnic	0.074	0.074	0.095	0.979
	(0.15)	(0.15)	(0.27)	
Covariates	No	Yes	No	
Observations	3924	3924	1020	

Notes: Ordered Logit specification, with dependent variable 0=not chosen, 1=indifferent, 2=chosen. Pooled data from the Non-election Round (July-August 2012) and Election Round (January-February 2013). The sample contains 981 individuals. Standard errors in parentheses are clustered at the individual level. P-values: * p<0.10, ** p<0.05, *** p<0.01. All specifications include fixed effects for each Dictator-profile (12 profiles in total). Covariates include interaction terms of the "Coethnic Profile" indicator with a gender indicator, education controls, and the Raven's test score. FWER p-values are simulated as described in the pre-analysis plan, for column (3). All specifications exclude ethnic Kamba subjects, as specified in the pre-analysis plan.

Table 5: Priming Effects in the Standard Dictator Game and Public-good Game

	Dictator Game Transfer (Percent of Endowment)		FWER p-value	Public-Good Game Contribution (Percent of Endowment)		FWER p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Election Round	-6.08***	-7.60***	0.008	-2.61	-6.36*	0.321
	(1.18)	(2.38)		(1.61)	(3.24)	
Political Competition Prime	-0.63	-2.19	0.591	0.26	-1.72	0.826
	(1.67)	(2.37)		(2.28)	(3.23)	
Ethnic-Cultural Prime	-1.45	-0.52	0.8	1.48	0.32	0.919
	(1.67)	(2.37)		(2.28)	(3.23)	
National Prime	-2.02	-4.43*	0.166	-0.37	-4.70	0.462
	(1.66)	(2.37)		(2.28)	(3.23)	
Election * Political Competition Prime		3.14	0.675		3.96	0.664
		(3.36)			(4.58)	
Election * Ethnic-Cultural Prime		-1.88	0.815		2.29	0.84
		(3.35)			(4.57)	
Election * National Prime		4.82	0.408		8.68*	0.321
		(3.35)			(4.56)	
Election * Ethnic-Political Prime		0.37	0.994		-1.77	0.964
		(2.38)			(3.24)	
Observations	1211	1362		1211	1362	

Notes: The dependent variables are the transfer in the Dictator game (in percent of the endowment) in columns 1-2, and the contribution in the Public-good game (in percent of the endowment) in columns 4-5. Pooled data from the Non-election Round (July-August 2012) and Election Round (January-February 2013). Standard errors in parentheses are clustered at the individual level. P-values: * p<0.10, ** p<0.05, *** p<0.01. FWER p-values are simulated as described in the pre-analysis plan, for columns 2 and 5. The Ethnic-Political priming, which was only implemented in the Election Round, is not included in columns 1 and 4. Each individual was endowed with 50 Ksh (approx. 0.60 USD) in the Dictator game, and with 60 Ksh (approx. 0.70 USD) in the Public-good game. All specifications exclude ethnic Kamba subjects, as specified in the pre-analysis plan.

Table 6: Coethnic Bias in the Implicit Association Test (IAT)

	Ethnic IAT: D-score				FWER p-values
	(1)	(2)	(3)	(4)	(5)
Constant	0.079***	0.013			
	(0.027)	(0.058)			
Political Competition Prime		0.046	0.15**	0.15**	0.133
		(0.081)	(0.067)	(0.066)	
Ethnic-Cultural Prime		0.077	0.11	0.082	0.459
		(0.085)	(0.067)	(0.066)	
Ethnic-Political Prime		0.100	0.10	0.088	0.459
		(0.086)	(0.067)	(0.066)	
National Prime		0.12	0.11	0.091	0.459
		(0.088)	(0.067)	(0.066)	
Order Controls	No	No	Yes	Yes	
Covariates	No	No	No	Yes	
Observations	547	547	547	547	

Notes: The dependent variable is the D-score, namely, the difference (in standard deviation units normalized by respondent) between the average response times in the Dissonance and Congruence IAT rounds. Data from the Election Round (January-February 2013). P-values: * p<0.10, ** p<0.05, *** p<0.01. FWER p-values are simulated as described in the pre-analysis plan, for column (3). Order Controls are indicators for the randomized order of (i) the ethnic and national IAT, and (ii) the Dissonance and Congruence rounds within each IAT; columns 1 and 2 also adjust for the different sampling weights of these randomized IAT order-groups. Covariates include ethnicity indicators, a gender indicator, education controls, and the Raven's test score. All specifications exclude ethnic Kamba subjects, as specified in the pre-analysis plan. Slow-response observations are excluded, as specified in the pre-analysis plan.