

The Politics of Language, Race, and Class:  
Experimental Evidence from South Africa

Thad Dunning

Department of Political Science

Yale University

This draft: May 10, 2010

Acknowledgements: Lauren Harrison did an extraordinary job of managing our research team in South Africa. Outstanding research assistance from Sadiyah Martin (project coordinator in Cape Town) and Gary Ramaru (project coordinator in Johannesburg), as well as Nadia Essop, Shumays Isaacs, Nayte Ngwenya, Bongani Ngubuka, Thulani Sibeko, Thozamile Ngqeza, Nick Rebelo, Thandi Mackenzie, and the rest of our team is gratefully acknowledged. I am also thankful to Jeremy Seekings and Ian Shapiro for their helpful advice. The staff at the Stats SA office in Pretoria graciously supplied data, while Yale's MacMillan Center for International and Area Studies and from the Institution for Social and Policy Studies provided financial support. Yale's Human Subjects Committee approved this research.

### Abstract

Much scholarship on South Africa has rightly focused on the politics of race and social class. Yet ethno-linguistic cleavages also played an important role in the transition to democracy, and they arguably shape the character of electoral competition today. To assess the relative weight of racial, class, and linguistic ties in shaping political preferences, I conducted an experiment with a probability sample of residents of Johannesburg and Cape Town, in which participants were shown similar videotaped political speeches by actors of varying race, native language, and stated class backgrounds. On average, I find strong race effects, some evidence for language effects, and perhaps counterintuitive class effects. Most interesting are the results for distinct sub-groups. There is strong evidence of a race effect for Whites and none for Blacks, while language effects appear to be stronger for linguistic groups that have arguably wielded less political power in the post-apartheid period. Finally, poorer subjects tend to favor richer candidates, an effect that is most pronounced for White subjects.

## 1. Introduction

Scholarship on South Africa has focused centrally on the politics of race and, to a lesser degree, social class to understand the character of electoral competition in the post-apartheid era. This is with good reason: apartheid was the dominant institution in South African life until the transition to democracy in 1994, and it was based explicitly on the racial cleavage. And while a strong racial identity helped to bridge class differences among whites (Lieberman 2003), social class also played an important secondary role in the transition to democracy (Wood 2000). Both race and class have also continued to shape politics since the transition. According to several accounts, the African National Congress (ANC) has succeeded in consolidating one-party rule by keeping alive the party's status as the bringer of liberation from apartheid, while debates about affirmative action programs such as Black Economic Empowerment have focused attention on the interaction between race and social class in contemporary South Africa.<sup>1</sup>

Yet, ethno-linguistic cleavages have also played an important role during and after the apartheid period in South Africa. The African National Congress has viewed itself as a pan-linguistic as well as pan-racial organization. Yet as leaders of Mangosuthu Buthelezi's Inkatha Freedom Party (IFP) pointed out during the transition, many of the party's key leaders were ethnically Xhosa—Nelson Mandela and Thabo Mbeki foremost among them—while Zulus and members of other Black African<sup>2</sup> ethno-linguistic groups were not as prominent. Indeed, Zulu nationalist violence incited by the IFP threatened to

---

<sup>1</sup> On social class, inequality, and affirmative action, see Seekings (2011) or Butler (2011).

<sup>2</sup> In this paper, I generally use the racial categories employed by Stats SA (the South African census office), which date from the apartheid era—that is, White, Black African, Coloured, and Indian/Asian—though I sometimes use Black in place of Black African.

destabilize the transition to democracy. Among White South Africans, meanwhile, it was not a foregone conclusion that conflict between English and Afrikaans speakers would take a back seat to the consolidation of the predominant white identity described by Lieberman (2003). The construction of the British colonial state in South Africa was predicated on bloody conflicts between British troops and Dutch settlers in the two Boer Wars (1880–1881 and 1899–1902), and the tension between the Anglo and Afrikaans communities was an important feature of social and political life well into the apartheid period (1948–1994). The history of confrontation and conflict along linguistic lines in South Africa pales perhaps only in comparison to the virulence of racial conflict.

Language politics arguably continues to influence the nature of democratic competition in South Africa today. The recent split in the leadership of the ANC appeared to have linguistic undertones, as supporters of the ousted Xhosa president Thabo Mbeki formed a new party, COPE, while leaders remaining in the ANC rallied around the new Zulu president, Jacob Zuma. Language politics have also played an important role in the Western Cape, where the Democratic Alliance (DA) party has won the only provincial premiership not held by the ANC. Winning support from Xhosa voters is crucial in the Western Cape, and the DA's White leader Helen Zille sometimes opens speeches with Xhosa phrases.<sup>3</sup> It also seems plausible that language politics could play an even bigger role in South Africa's future. Linguistic differences might even emerge as a new basis for political mobilization, in a way that could eventually undermine the hegemony of the ANC—much as the rise of regional linguistic parties in India helped to undermine the single-party dominance of the Congress Party.

---

<sup>3</sup> Zille speaks Xhosa along with English, Afrikaans, and German, the language of her parents.

For political scientists, the structure of linguistic and racial cleavages in South Africa—in which first language groups are largely, though not entirely, nested within racial groups—therefore raises interesting questions about the conditions under which a super-ordinate category like race remains a dominant cleavage (Dunning 2009)—and about the conditions under which alternate linguistic or class categories could take on greater political importance. Yet, assessing the role of language cleavages in shaping political attitudes and behavior as well as broader political outcomes is made more difficult by the very electoral hegemony of the ANC. Despite the rise of the Democratic Alliance in the Western Cape and the emergence of COPE during the most recent elections, the ANC remains the dominant party in South Africa, and it attracts support from Black African voters of various linguistic origins. Aggregate electoral returns thus give us little insight into the actual or potential political role of language. Nor does existing data readily allow assessment of the independent role of social class in shaping electoral behavior.

To assess the relative weight of racial, class, and linguistic ties in shaping voter preferences, I therefore conducted an experiment with a probability sample of residents of Johannesburg and Cape Town. Participants were shown identical videotaped speeches in which the race, native language, and stated class background of an actor posing as a political candidate were varied. The experimental design is described in greater detail below.

The experimental evidence suggests that both race and language matter: on average, I find strong race effects and some evidence for language effects. Yet there is intriguing heterogeneity in these effects across racial and linguistic sub-groups. For

example, while I find race effects for White participants in the experiment—who are significantly less likely to say they will vote for the candidate when exposed to a Black actor than when exposed to a White actor—I do not find such race effects among Black voters. Language effects may be strongest for minority groups that have arguably been relatively marginalized politically in the post-apartheid order, such as Afrikaans- and Zulu-speakers. Finally and unexpectedly, poorer subjects tend to favor richer candidates, an effect that is most pronounced for White subjects. Below I discuss several aspects of the racial, class, and linguistic cleavages in South Africa and then describe the experimental design, before turning to the data analysis.

## **2. Race, Language, and Class Cleavages in South Africa**

With 11 official national languages,<sup>4</sup> and a complex system of racial classification inherited from the apartheid era, South Africa would seem to be a setting with a quite complex social cleavage structure.<sup>5</sup> Yet several observations help to clarify the structural relationship between race and language and set the stage for the experimental study described in the next section. First, native language is typically nested within race in South Africa, so that White South Africans nearly always speak Afrikaans or English as a first language, while so-called Black Africans speak Zulu, Xhosa, or one of South Africa’s other official or unofficial languages. This “nested” cleavage structure is of general interest to political scientists, who may seek to understand the conditions under

---

<sup>4</sup> They are Afrikaans, English, IsiNdebele, IsiXhosa, IsiZulu, Sepedi, Sesotho, Setswana, SiSwati, Tshivenda, and Xitsonga. In this paper, I refer to Xhosa and Zulu rather than IsiXhosa and IsiZulu.

<sup>5</sup> Here I use the term “social cleavage structure” simply to refer to the joint distribution of membership in different racial and linguistic as well as class groups.

which a super-ordinate or subordinate ethnic identity category becomes politically salient (Posner 2005; Dunning 2009).

Second, while a number of Black African language groups are demographically and politically relevant in post-apartheid South Africa, the dominant Black African language groups in the political arena have clearly been the Xhosa and the Zulu. This is not to say that native speakers of other languages—perhaps most notably the Sotho—have not played important political roles. It is simply to emphasize that understanding the interaction of Zulu and Xhosa linguistic cleavages is of particular interest, and this will justify the focus on these groups (along with English and Afrikaans) in the experimental study below.

Third, notwithstanding this principle inquiry into the salience of the Black and White racial categories and of the linguistic groups nested within them—that is, groups of Zulu- and Xhosa-speakers among Blacks, and English- and Afrikaans-speakers among Whites—it also appears important to understand the relationship of the Coloured category to these other racial and linguistic categories. Survey data described below suggest that while the majority of Coloured—located principally in the Western Cape but also in the Eastern Cape, Natal, and the Northern Cape—speak Afrikaans as a native language, a large proportion speak English as a native language, making the relationship of language politics to this racial category of particular interest. Although this paper primarily describes an experiment involving Black African and White subjects and politicians, experimental subjects were drawn from a probability sample that includes a large proportion of Coloureds, and future work will incorporate analysis of data on the effects of politicians' language, race, and class on the candidate preferences of this group.

Finally, the relationship of social class to these other cleavages is also of principal interest, in the experiment to which I turn next. For much of the apartheid period, it is safe to say that race and class formed overlapping rather than cross-cutting cleavages (Taylor and Rae 1979), with even poor Whites being substantially better off than the best-off Blacks.<sup>6</sup> In the post-apartheid period, the extent to which this is true has lessened, with the political empowerment of the ANC and the parallel growth of economic power for select Blacks in the private sector. The relationship between race and class, and the role of racial and class ties in shaping political preferences, is thus of special interest in contemporary period.

### **3. Experimental Design**

How, then, do racial, ethno-linguistic, and class ties shape preferences for political candidates in South Africa? I developed an experiment that allows comparison of the relative weight of these three kinds of identities in shaping voter preferences, using an experimental design similar to one I have used in my research in Mali, India, and Brazil (Dunning and Harrison 2010; Dunning 2009). In brief, subjects were exposed to videotaped speeches by actors posing as political candidates; the racial, linguistic, and class relationship of the subject to the actor/candidate was varied experimentally, in ways described below. The experimental design aimed to isolate the effects of racial, linguistic, and class ties on subjects' evaluations of candidates.

In total, 755 Black African and White experimental subjects—211 native Zulu speakers and 198 native Xhosa speakers among the former and 239 native English

---

<sup>6</sup> Of course, this is not to deny the important heterogeneity in wealth and income within the White group (Lieberman 2003).



speakers and 107 native Afrikaans speakers among the latter—were recruited via a probability sample of South African residents of Johannesburg and Cape Town.<sup>7</sup> These subjects were exposed at random to videotaped speeches, given by Black African and White actors posing as potential candidates for the National Assembly (the lower house of Parliament); the candidate’s political party was not mentioned.<sup>8</sup> Each actor recorded two versions of a speech designed to mimic typical political speeches given during campaigns for the National Assembly. In the first version, the actor professed a “popular” working-class background, wore ordinary street clothes, and expounded a message focused on popular issues such as inequality, unemployment, housing, access to basic education, and poverty; in the second version, the actor professed a more elite occupational and educational background, wore a business suit, and expounded a more elite message focused on reducing regulation, promoting private clinics, bolstering opportunities for the middle-class and promoting merit-based recruitment.<sup>9</sup> The speeches were mainly given in English, but the opening and closing phrases of the speeches were in Zulu, Xhosa, English, or Afrikaans, according to the native language of the actor in the video. Subjects were shown these speeches on portable DVD players or laptops, as part of the door-to-door survey described below, and then were asked a battery of questions tapping the degree to which they would want to vote for the videotaped candidate as well as the candidate’s perceived competence, likeability, and so on.

---

<sup>7</sup> As described below, these subjects were recruited as part of a much broader probability sample that included members of other racial groups, including Coloureds and Indian/Asians, and Black Africans from other linguistic groups. I restrict much of the analysis to Zulu- and Xhosa-speaking Black Africans and Whites for reasons discussed elsewhere.

<sup>8</sup> The advantages of failing to prime a party affiliation are mentioned below.

<sup>9</sup> The speeches are described in more detail below. The full text of the speeches, along with the post-treatment questionnaire, are available at <http://www.thaddunning.com/data/south-africa>.

Subjects were thus exposed at random to a speech by an actor/candidate who (1) shared or did not share their race (Black African or White); (2) shared or did not share their native language (Zulu, Xhosa, English, or Afrikaans); (3) shared or did not share their same broad income group (“rich” or “poor”). Because I presumed that native language is nested inside race when designing the experiment—that is, White subjects have a native language of English or Afrikaans, while Black African subjects in this experiment have a native language of Zulu or Xhosa<sup>10</sup>—the crossing of race and language generated three treatment conditions: (i) subject and politician have the same race and native language; (ii) subject and politician have the same race but different native languages; and (iii) subject and politician have different races and native languages. (The assumption that language is nested within race here is borne out by my survey data.<sup>11</sup>) Within each of these three conditions, subjects could be exposed to a politician from their own income group (rich or poor) or from a different income group; the operationalization of these categories is discussed below.

This set-up thus generated the experimental design depicted in Table 1, which shows the number of subjects assigned to each treatment condition. The table reports a cross-tabulation of the race and language treatments, with the social class treatments reported inside each cell. Notice that within each cell, the sum of subjects assigned to the “same income” and the “different income” conditions does not total the number of

---

<sup>10</sup> Data for Black African native speakers of other languages—Sotho, Tswana, and so on—is analyzed below, but like Coloured and Indian/Asian respondents, these respondents could not be assigned to one of the experimental conditions in Table 1.

<sup>11</sup> There were no White subjects who identified Zulu or Xhosa as their first language, while among Black Africans, 14 subjects said their first language is Afrikaans and 14 said their first language is English. For only 19 of these 28 Black African subjects did the investigator identify English or Afrikaans as the subject’s first language after completing the interview, however; for 2 of those subjects, the interview actually took place in a different African language (Sesotho).

subjects assigned to the race/language condition given by that cell; this is because of missing data on the income question. When I analyze the effects of the social class treatments, rather than pooling across them, the size of the experimental study group is therefore reduced from 755 to 688. However, the presence of missing data should in principle be independent of the randomized treatment assignment, and analysis suggests that missingness and treatment assignment are indeed uncorrelated at the level of the data.

[TABLE 1 ABOUT HERE]

One important point to note about the experimental design is that here voters are asked evaluate individual candidates. It might be pointed out that South Africa has a list system with proportional representation, which might seem to make the evaluation of individual candidates less politically relevant. Yet following the logic of Carey and Shugart (1995), such single party-dominant systems may nonetheless create strong incentives for politicians to cultivate the personal vote, for instance, as a means of moving up in the order of party lists. Moreover, attitudes towards candidates that are captured in the experiment give clues towards a broader constellation of racial, linguistic, or class effects with potentially larger-scale consequences for the political system as a whole.<sup>12</sup>

### *Manipulating perceptions of race and language*

I now turn to a fuller discussion of how perceptions of candidates' race, native language, and social class were stimulated. First, with respect to race, subjects were

---

<sup>12</sup> This is also a justification for not having candidates mention political party in their speech; here, we hope to elicit the effects of race, class, or native language on subjects' evaluations, free of association with a party such as the ANC.

randomly assigned to view speeches by actors/candidates who were either Black African or White. Though, as discussed below, this strategy raises some challenging issues for purposes of making causal inferences, it is obviously a conceptually straightforward means of stimulating perceptions of race in South Africa.<sup>13</sup>

Manipulating perceptions of native language involved a slightly subtler strategy. First, the introduction and the closing sentences of the speech were delivered in the actor's native language, that is, in English, Afrikaans, Zulu, or Xhosa. The actor began the speech by saying (the text in *italics* was spoken in the candidate's native language, while the subsequent text was in English for all candidates):

*I greet you all, and thank you for being here today. I am [CANDIDATE'S NAME] and it is an honor to be able to speak to you about South Africa, and about issues that matter to all of us. My friends, I am here today to ask you for your support in my candidacy for the National Assembly...*<sup>14</sup>

Next, the given name and surname of the candidate were varied according to his intended ethnicity. In the text above, CANDIDATE'S NAME is substituted with Daniel Mitchell for the English speaker, Danie Meiring for the Afrikaans speaker, Dingane Mazibuko for the Zulu speaker, and Umzileapheli Botha for the Xhosa speaker—all names that seemed to local informants to indicate the ethnic identities we intended. Finally, because actors were recruited to “play the part” of a politician from their own native language group, any differences in accent across language groups would also show up in the English-language portions of the speech.<sup>15</sup>

---

<sup>13</sup> Over 90 percent of subjects successfully identified the race of the politician, when responding to a question asked after the experiment.

<sup>14</sup> The politician also returned to his native language for the speech's closing sentence: “I thank you for your attention and for your support, and look forward to working on your behalf.”

<sup>15</sup> Around 70 percent of subjects correctly identified the first language of the politician, with the highest correct proportion (over 80 percent) for English-speaking politicians, then Afrikaans-speaking politicians, and finally for Zulu and Xhosa politicians. Some of the high error rate for

Though these are fairly straightforward means of stimulating (and manipulating) perceptions of a candidate's race and native language, some subtleties present themselves: drawing causal inferences about the *effects* of perceptions of race on candidate evaluations is not as straightforward. Indeed, a major challenge in the analysis of race and language effects is that race and first language cannot, in fact, be experimentally manipulated. Imagine asking White subjects to compare two videotaped political candidates, one of whom is White and the other of whom is Black, and suppose that White subjects on average rate the White candidate more favorably than they do the Black candidate. This difference would not necessarily provide *prima facie* evidence of stronger in-group preferences (a.k.a. racism) on the part of White subjects: it could simply be that *this* particular White candidate is genuinely more attractive or compelling than *this* Black candidate, for race-independent reasons. (For example, the white candidate might happen to be more articulate, more charismatic, etc.). This poses difficulties for attempts to isolate the effects of racial or linguistic ties using an experiment: when we change an actor's (candidate's) race, we also change many other things about him (Holland 1986). With one White and one Black actor, we cannot hold such actor fixed effects constant, and so the effects of race are not identifiable.<sup>16</sup> The same argument holds for language, since we cannot readily change the first language of an actor without changing many other things about him.<sup>17</sup>

---

the latter is due to White respondents who could not distinguish between the Xhosa and Zulu languages. However, for all candidates and types of subjects, the modal identified language is the intended language of the candidate.

<sup>16</sup> That is, distinct parameter values could give rise to the same joint distribution of the observables, leading to a problem of observational equivalence.

<sup>17</sup> We considered hiring perfectly bilingual actors and asking them to give the introduction to the speech in two languages (English and Afrikaans or Zulu and Xhosa); this would allow us to experimentally manipulate the language of the politician, at least within racial group. However,

We can appeal to two statistical arguments to (at least partially) elide this difficulty. First, if we could sample many White and Black actors at random, the race- or language-independent attractiveness of the candidates should be more or less balanced across the two groups, up to random error, with the balance getting better as the number of actors in each group grows.<sup>18</sup> Differences in average evaluations of White and Black candidates by White and Black subjects could then be more reliably attributed to race-of-candidate effects. Our procedure does not quite attain this optimum, since we could not realistically sample actors at random (defining the relevant population is another matter altogether). However, we did recruit White and Black actors of varying ages, education levels, and other attributes, and our pool of White and Black actors seems roughly comparable.<sup>19</sup> And while our sample of actors is not huge, for logistical as well as research-design reasons,<sup>20</sup> we ended up with twelve actors—six White and six Black—across which to make comparisons, which helps considerably with efforts to hold estimate the effects of race, net of race-independent candidate fixed effects.

Second, in our analyses we will frequently compare the *difference* in Black subjects' evaluations of White and Black candidates to the *difference* in White subjects' evaluations of White and Black candidates. Such difference-in-difference comparisons

---

this proved impractical. Even if we could have found actors with the ability to give a perfectly unaccented introduction in each of the respective languages, their accent in the English portion of the speech would tend to indicate their first language.

<sup>18</sup> Of course, this is not true of race-*dependent* characteristics; yet race-dependent perceptions of attractiveness are exactly what we want to study.

<sup>19</sup> We recruited actors by posting advertisements in local stores and through word-of-mouth; several actors, including Zulu, English, and Afrikaans speakers, were teachers at a private school in Johannesburg, while one Xhosa speaker had been a union organizer and is now a political candidate (not in Johannesburg or Cape Town, however). None of the candidates were professional actors or otherwise recognizable to subjects.

<sup>20</sup> There is a tradeoff in increasing the number of actors for a fixed experimental population size: as the number of actors grows, the number of subjects exposed to each actor diminishes, which limits the precision with which we can estimate within-actor effects (for example, of experimentally-manipulated attributes such as the candidate's social class; see below).

also help to allay concerns about candidate fixed effects. Even if we happened to recruit White candidates who are simply more attractive or charismatic than the Black candidates we recruited, if participants in our experiment do not have in-group racial preferences, then the difference in evaluations of White and Black candidates on the part of White subjects should be about the same as the difference in evaluations of White and Black candidates on the part of Black subjects. As I show below and in the Appendix, these statistical arguments will help us to separate true race and language effects from race- and language-independent attributes of the actors.

### *Manipulating perceptions of social class*

Perceived social class is more plausibly subject to true experimental manipulation, because the same actor can be used to stimulate different class treatments. In this experiment, the perceived social class of the videotaped actor/politician was stimulated in several ways. First and most basically, the dress of the actor varied across the “poor” and “rich” conditions; in the former, he wore a simple collared shirt or other more informal street wear, while in the latter, he wore a business suit. Secondly, the stated background of the candidate varies across two versions of the speech filmed by all of our actors. In the “poor” speech, the politician follows the introduction discussed above by saying:

I have been a working person all of my life, so I know the struggles that ordinary people face every day....I was born, raised and educated in the working class neighborhoods of South Africa. Jobs are not easy to come by, but I eventually found work, first as a taxi driver in Johannesburg and then as a mechanic in a small factory in the Western Cape. After several years in the factory, I grew frustrated by the conditions so I became involved in the local mechanics’ union. Since that time, I have been working both as a mechanic and as an organizer with the union, fighting for the rights of working people.

In the “rich” speech, this text is replaced with:

I have been a professional and an active community member all of my life...I graduated from the University of Cape Town and did my post-graduate education there as well. For the past fifteen years, I have worked as a lawyer in Johannesburg, and I have also been a business owner in the neighborhood of Sandton.<sup>21</sup> Throughout the course of my career, I have worked with many organizations to improve the quality of life in our neighborhoods and especially to improve the quality of schools in our cities.

The profession of the politician was also reinforced by showing a subtitle near the introduction to the speech, which showed the politician’s assigned name and profession—taxi driver and mechanic in the case of the working-class politician, lawyer and business owner in the case of the more elite politician.

Finally, we also varied the types of social and political problems on which “poor” and “rich” politicians focused. The speech of the poor politician focuses on jobs and poverty, noting that

our struggle is against poverty and social inequalities and in defense of the interests of the common people. Our struggle is to make life better for the average South African.

The speech of the rich politician, in contrast, dwells on the problems of business owners and university graduates. The rich politician promises to employ his

years of experience as a businessman and as a lawyer to fight for the middle-class and to expand educational and employment opportunities in South Africa...crack down on crime...[and] encourage innovation and entrepreneurship, not more government regulation and redistribution.<sup>22</sup>

---

<sup>21</sup> Sandton is an elite residential and business neighborhood in the suburbs of Johannesburg.

<sup>22</sup> The politician also makes an appeal to merit-based advancement, saying “I will defend a policy of merit in our schools and universities and in our labor markets and not a policy of special privileges. The law should not differentiate between citizens; it should protect and assist all South Africans. We should reward people who conquer spaces with their intelligence and ambition, not with government favors.”



Of course, one disadvantage of bundling all these various stimuli—the dress of the politician, his stated background, and the types of problems on which he would focus if elected—into a single “treatment” is that we will not readily be able to discern what aspect of the treatment has an effect. In particular, we cannot distinguish the effects of the politician’s “message” from other aspects of class background. However, the rationale for bundling these components in this experiment is that we want to compare the broad effects of socio-economic class to those of race and language, and so it is useful to start with a broad and hopefully effective means of stimulating perceptions of the politician’s class orientation. After the experiment was concluded, we asked subjects to locate the politician’s income group on a 5-point scale (we showed subjects a visual aid, in which the highest category had stacks of coins and dollar signs and the lowest category had a single coin).<sup>23</sup> Our data suggests that we were able to stimulate perceptions of the class background of the politician with some success.<sup>24</sup> Note that in the analysis below, we will often pool across the two versions of the speech (rich and poor), in order to focus attention on comparing the effects of race and language. Thus, the precision with which we might stimulate different aspects of a “class effect” is of somewhat secondary interest for present purposes.

---

<sup>23</sup> The question wording was as follows: “A recent survey by the Bureau of Market Research at the University of South Africa ranked the incomes of citizens as A, B, C, D, or E, with A corresponding to the richest people and E corresponding to the poorest. In your opinion, to which income group does this politician belong?”

<sup>24</sup> Rich politicians were placed about 0.3 points higher on the 5-point scale, or about a third of a standard deviation.

### 3.1 Sampling Participants

Before turning to the analysis, it is useful to describe the recruitment of survey respondents, i.e., participants in the experiment. The goals in developing a sampling strategy were two-fold. For purposes of increasing the precision of treatment effect estimators for different sub-groups of the population, it was desirable oversample relatively low-frequency population types (such as poor Whites and rich Blacks). Yet for reasons discussed above, it was also desirable to obtain a strict probability sample of the target population, to the extent possible, which implies that quota sampling is inappropriate. Obtaining a strict probability sample was desirable because it allows us to extrapolate responses to (non-experimental) survey responses, as well as estimates of treatment effects, to the target population. Further, while the experimental analysis below focuses mainly on the responses of (English- and Afrikaans-speaking) Whites and (Zulu- and Xhosa-speaking) Black Africans, it will also be of interest to analyze the candidate evaluations of other respondents—such as Coloureds—who showed up with great frequency in our probability sample.

Because we lacked a sampling frame that identified individuals' class, race, or first language, our strategy was to oversample respondents from neighborhoods—called “sub-places” by the South Africa statistics agency, Stats SA—with certain aggregate demographic characteristics, e.g., high proportions of poor Whites and rich Blacks.<sup>25</sup> First, we received from Stats SA data on the populations of 472 sub-places (all 314 in metropolitan Johannesburg and all 158 in metropolitan Cape Town); these sub-places

---

<sup>25</sup> Cross-tabulations by race, income, and first language at the sub-place level are based on census data and were provided to us by Stats SA.

range in size from 3 people<sup>26</sup> to 55,330 people, with an average population of 3,975 and a median of 1,916. Next, we sampled sub-places at random with probability proportionate to population; the numbers of interviews to conduct in each sub-place were determined in proportion to the population of the selected sub-places. Finally, we added a fixed number of additional interviews in sub-places selected according to the above-noted demographic criteria; this over-sample served the objective of making more likely the adequate representation of low-frequency population types.<sup>27</sup> Note that since we have population data by sub-place and know how many respondents were selected from each sub-place, we may determine the ex-ante probability of selection for each respondent in the sample. Thus, we can use post-sampling weights (weighting respondents by the inverse of their probability of selection) in our analysis to produce parameter estimates that are valid for the whole universe from which the sample was drawn.

After selecting sub-places and determining the number of respondents to be sampled from each sub-place, we sampled respondents at random from within sub-places, using detailed maps of each sub-place obtained from Stats SA. These maps allowed us to delimit street corners and intersections and select these at random as starting places for our interviewers. We also used a random method to draw walking routes on the maps for our enumerators, who then used an interval sampling method to select households and the method of birthdays to select individual respondents from within households.<sup>28</sup> While

---

<sup>26</sup> A few sub-places are comprised of institutions such as prisons, and prisoners are apparently not counted as residents of the sub-place by Stats SA.

<sup>27</sup> We oversampled respondents from the following sub-places. *Johannesburg*—Berea, Eldorado Park, Hillbrow, Joubert Park, The Hill. *Cape Town*—Acacia Park, Airport Informal, Bishops court, Durheim, Muizenberg, Oranjezicht, Signal Hill/Lions Head, Westlake, Wittenboomen, Zonnebloem.

<sup>28</sup> The person who answered the door was asked to list all members of the household and their dates of birth; the person with the nearest birth date to the date of the interview was selected. If

calculating the real response rate is a somewhat imprecise task, we estimate that it was about 33 percent.<sup>29</sup>

This sampling strategy did produce a sample with more relatively rare types, such as poor Whites and rich Blacks, than would have a simple random sample or other self-weighting design (Table 2).<sup>30</sup> The main experimental analyses include a group of 633 subjects that is comprised of 18.2 percent (N=115) rich White English-speakers and 10.1 percent (N=64) poor White English-speakers, where “rich” means that the respondent placed him or herself in the top 3 of 9 income categories on an ordinal scale. (Using an educational proxy for income—whether the respondent graduated from college or university—or a coding based on investigator perceptions produces a similar distribution

---

that person was not available, investigators asked for the person with the next-nearest date of birth, counting the missed interview as a refusal for purposes of calculating the response rate. Domestic helpers and other workers were not included in the list, as we hoped to sample such people in their own households.

<sup>29</sup> The real response rate is calculated as follows. First, we asked interviewers to record the number of houses they approached and whether an interview was conducted in each house or not. Based on these records, we estimate that interviews were conducted successfully in about 43.12% of households approached. Next, we also asked interviewers to note the number of unavailable respondents selected within households using the birthday method, in each house in which they successfully conducted an interview. In 845 of the 1,520 households for which these data were recorded, the first selected respondent was interviewed; in 375, the second was interviewed; in 201, the third; in 68, the fourth; in 21, the fifth; in 5, the sixth; and in 5 households, the seventh selected respondent was actually interviewed. This implies that a total of  $(375)(1) + (201)(2) + 68(3) + (21)(4) + (5)(5) + (5)(6) = 1,120$  refusals were needed to obtain 1,520 completed interviews, which gives a ratio of  $1,120/1,520=0.7368$  refusals per successful interview.

To obtain the 1,892 interviews in our database, we can estimate that a total of  $1,892/0.4312 = 4,388$  households were approached and that within the 1,892 households in which an interview took place, there were an additional  $(0.7368)(1,892)=1,394$  refusals. That gives 5,782 attempted interviews—that is, 1,892 successful interviews, 1,394 refusals within households where interviews were successfully completed with other respondents, and 4,388 refusals from households where no interview took place.

Thus, we have an estimated response rate of  $1,892/5,782$ , or about 33 percent. This may slightly overestimate the true response rate, since we could not obtain records on the number of houses approached from every field investigator (and the ones who did not turn in records may conceivably have had lower response rates).

<sup>30</sup> Table 2 codes participants as “rich” if they identify themselves as in the top 3 of 9 income categories. Alternate composite measure that also codes respondents as rich if they have completed college or university, or if the interviewer places them in the top 2 of 5 income groups, produce similar distributions as in Table 2.

as in Table 2, as does the use of investigator perceptions of the race and first language of the respondent). This is certainly a much higher proportion that we would have obtained through a self-weighting design. Also, a relatively high 7.4 percent (N=47) and 8.4 percent (N=53) of the experimental study group is comprised of rich Black Zulu-Speakers and rich Black Xhosa-Speakers, respectively. On the other hand, just 4.4 percent (N=28) and 5.8 percent (N=37) of the study group is poor White Afrikaans-Speakers and rich White Afrikaans-Speakers, respectively. Poor Zulus (N=155) and poor Xhosas (N=134) together comprise 45.7 percent (N=289) of the experimental study group, while Blacks as a whole comprise 61.5 percent (N=389) of the participants in the experiment.

[TABLE 2 ABOUT HERE]

Yet, the sampling strategy had some costs as well. The numbers of Afrikaans speakers, in particular, is quite low, which limits the precision with which race, class, and language effects may be estimated for this sub-group. And the lack of an individual sampling frame and the desire to draw a probability sample implied an important degree of inefficiency, since 960 respondents out of 1,892, or nearly 51 percent, were Coloured (N=494)<sup>31</sup>, Indian/Asian (N=78), Other (N=87), or were Black native speakers of languages other than Zulu and Xhosa (N=312).<sup>32</sup> Another 299 respondents did not identify their income, first language, or race. Like the rest of the respondents, these respondents were exposed to the various experimental stimuli at random, and I analyze their responses to these stimuli below. Yet since none of them can be assigned to all of

---

<sup>31</sup> This high incidence of Coloureds reflects the choice to work in the Western Cape: 392, or nearly 80 percent, of the Coloured respondents resided in Cape Town.

<sup>32</sup> Among Blacks, the major non-Zulu and non-Xhosa first language was Sotho (N=130), with other respondents listing English (N=43), Afrikaans (N=35), Tswana (N=31), Swati (N=24), Tsonga (N=16), Venda (N=17), and Ndebele (N=13).

the main treatment conditions listed in Table 1—because by design none of our actors were Coloured, so these respondents cannot be assigned to conditions in which they share the race the politician/actor—such Coloured, Indian/Asian, and non-Zulu/non-Xhosa Blacks are excluded from the main experimental analyses reported below. However, these data will be analyzed in a future version of this paper.

#### **4. The Effects of Race, Language, and Class**

After being shown the videotaped speech, respondents were shown a picture of a ladder and asked the following question: “Please look at this ladder, which has seven steps. Suppose the first step of the ladder means ‘no, not at all’ and the 7th step means ‘yes, completely.’ Where would you put your answer to the following question: Does the speech make you want to vote for this candidate?” Answers to this and other questions about the candidate and the speech, pooling across treatment conditions, are presented in the table of descriptive statistics (Table 3).

[TABLE 3 ABOUT HERE]

What are the effects of racial and linguistic ties on candidate evaluations? Table 4 presents the average responses to this question, by assignment to three treatment conditions: (1) the subject and candidate have the same race and same native language; (2) the subject and candidate have the same race and different native languages; (3) the subject and candidate have different races and languages. The analysis in Table 4 pools across the two social class treatments (which were assigned at random with equal probability within each cell of the table), focusing on the causal effects of treatment assignment along the race and language dimensions. The difference between any two of

(1), (2), and (3) estimates the average causal effect of treatment assignment—that is, intention-to-treat parameters.<sup>33</sup>

[TABLE 4 ABOUT HERE]

This intention-to-treat analysis suggests the presence of race effects as well, perhaps, as more modest language effects. The difference in average candidate evaluations of subjects assigned to the same-race-same-language condition and the different-race-different-language condition is 0.51 points on the seven-point scale, a highly significant effect of nearly one-third of one standard deviation. Clearly, on average, race and language together shape subjects' evaluations of the candidates.

What are the separate contributions of race and language to this total effect? The analysis in Table 4 suggests that on average, race is slightly more important. The estimated effect of exposure to a politician who speaks the same native language, relative to one who speaks a different language but has the same race, is 0.21 points; with a standard error of 0.18, this estimate is not statistically significant. The difference in evaluations of a politician of the same race—but different language—and a politician of a different race and native language is 0.30—and with a standard error (0.16), this estimate is significant at the 0.1 level. Thus, while these point estimates are not very different, there is some evidence that the contribution of race is, on average, stronger than the contribution of language.

---

<sup>33</sup> To analyze Table 3 and the other results, I appeal to the Neyman (1923) model, also known as the Neyman-Rubin-Holland model. Here, subjects have potential outcomes, which are their evaluations of each politician under each treatment assignment. The average causal effect for any two treatments is the difference between average evaluations, if all subjects were assigned to the first treatment, minus average evaluations if all subjects were assigned to the second treatment. Because of random assignment, a simple difference-of-means—the difference between average evaluations of subjects assigned to the first treatment and those assigned to the second treatment—suffices to estimate this parameter.

### *The contrasting effects of race*

Yet, the analysis in Table 4 masks substantial and quite interesting heterogeneity in the effects of treatments. Perhaps most striking is the total absence of a race effect for Black subjects—and the presence of a strong race effect for White subjects. As Table 5 shows, while the difference between evaluations of same-race and different-race politicians is a highly significant 0.55 for White subjects, this difference is estimated at 0.00 for Black subjects.<sup>34</sup> In other words, there is no evidence of in-group preferences for Blacks, but strong evidence of in-group preferences for Whites. (We find below that this pattern also holds for each individual actor: White subjects evaluate each individual White candidates more positively and each individual Black candidates less positively, while for Blacks, there is no systematic difference).

[TABLE 5 ABOUT HERE]

### *The effects of language*

How do language differences shape subjects' evaluations of candidates? Table 6 compares average evaluations of candidates who share the subject's language and those who do not, for subjects from each of the four linguistic sub-groups (Afrikaans, English, Xhosa, and Zulu). Note that this comparison in some ways conflates the effects of race and language, because politicians who share the subject's language (and thus race) are here compared to those who do not; the latter group of politicians may have the same race

---

<sup>34</sup> Notice that while the sample size is more than half as small for the White-only group as for the whole experimental population, the standard error on the difference of 0.55 for Whites is just as small as the corresponding standard errors in Table 3. This is because the standard deviation of candidate evaluations (within treatment assignment categories) is substantially smaller for Whites than for blacks.



as the subject or they may have a different race. Still, this is a useful first cut at the issue. In particular, Table 6 suggests interesting heterogeneity across sub-groups. Here, there are apparently strong language effects for English-speakers (though see below), and strikingly, much stronger effects for Zulu speakers than for Xhosa speakers. Indeed, there is little difference among Xhosa speakers in evaluations of politicians who share their language and those who do not, while for Zulu speakers there is a difference of 0.57 points, an estimated effect that is significant at the 0.1 level.

[TABLE 6 ABOUT HERE]

A better comparison for purposes of assessing language effects, however, may be within racial groups. For instance, among Whites, do Afrikaans-speaking subjects prefer Afrikaans-speaking politicians to English-speaking politicians? Among Black Africans, do Zulu-speaking subjects prefer Zulu-speaking candidates to Xhosa-speaking candidates?

To answer these questions, Table 7 compares the average evaluations of same-race, same-language politicians to average evaluations of same-race, different-language politicians, for the Afrikaans, English, Zulu, and Xhosa sub-groups. Unfortunately, this stratification results in quite small sample sizes and large standard errors; here, within-group estimates are not significant. Yet the point estimates are large, and for Afrikaans- and Zulu-speakers, they exceed by far the average effects of language estimated above (see Table 4). On the other hand, for English and Xhosa speakers they are much smaller. We comment on these and other results further below.

[TABLE 7 ABOUT HERE]

Notice that the result for English-speakers in Table 7 appears to contradict Table 6. This suggests that what is driving the apparent effect in Table 6 is race, rather than language. Indeed, when we compare evaluations among White English-speaking subjects of White politicians (both English- and Afrikaans-speaking) and Black African politicians (both Zulu- and Xhosa-speaking), we find a strong preference for White politicians—the difference of 0.64 points is highly significant. This difference is dwarfed only by the comparison among White Afrikaans speakers, who on average rate White politicians a whopping 1.84 points higher than Black politicians—a difference of nearly one standard deviation that is significant at the 0.001 level.

It is perhaps suggestive that we find the strongest race effects among the politically-disempowered minority racial group—Whites—yet also find language effects among the language group that is arguably the relatively disempowered group, within each racial category.<sup>35</sup> Indeed, this finding could be consistent with theories of ethnic politics in which the political empowerment of a superordinate category—in contemporary South Africa, Blacks—generates greater competition among groups comprising subordinate categories, here, Zulu and Xhosa. While this may reflect a more general logic (Posner 2005), the extent to which greater competition between members of the subordinate category may be engendered, rather than greater solidarity, may vary (Dunning 2009). Future work will be oriented towards better understanding and explaining the effects uncovered in the analysis of race and linguistic cleavages in this section.

---

<sup>35</sup> Post-treatment survey questions tapped respondents' perceptions of the most and least politically-powerful linguistic groups in South Africa; a future version of this paper will analyze those data.

### *The effects of social class*

What about social class effects? Recall that subjects were randomly assigned to view a speech by a “rich” or a “poor” politician, where the candidate’s dress, stated background, and message were used to manipulate perceptions of the social class of the politicians. Next, subjects can themselves be classified as “rich” or “poor,” though one has to decide how to do the classification; in the analysis here, I code those who self-identified their income in the top 3 of 9 categories as rich, coding the rest as poor. Thus, one can compare evaluations of politicians from subjects’ own income groups—rich or poor respectively—to evaluations of politicians from the opposite income group. This provides an initial means of looking at the salience of perceived income in shaping stated voting preference.

The results in Table 8 are interesting, and apparently counterintuitive. Politicians’ from subjects’ own income group are rated significantly *less* positively, on average: the difference is 0.23 points, which is significant at the 0.05 level. Sub-group analysis suggests that the evaluations of poor respondents, and especially poor Whites, drive this finding. Among all poor subjects, rich politicians are preferred to poor politicians by 0.34 points (significant at the 0.01 level), while among poor Whites, rich politicians are preferred to poor politicians by 1.10 points (also significant at the 0.01 level).<sup>36</sup> For rich subjects, in contrast, the point estimate suggests that rich politicians are rated 0.07 points higher than poor politicians, but this difference is not statistically different from zero.

There is some evidence that rich White subjects do prefer rich politicians, but no

---

<sup>36</sup> In Table 8, I do not adjust the reported significance levels for multiple comparisons, though it would be straightforward to do so; the main contrast in the table is between effects for rich and poor subjects, while the additional data provides for additional texture.

evidence that rich Blacks do so. In sum, the preference for rich politicians is much more pronounced among poor subjects.

[TABLE 8 ABOUT HERE]

Of course, the best experimental comparison may be *within* racial and linguistic treatments: in this case, only the text of the speech and the dress of the actor varies, while attributes of the actor himself are held constant. I find very similar effects here (full results omitted). For example, among poor Whites of either language group who are exposed to a speech by a White candidate also of either language group—that is, holding the racial (but not linguistic) treatment constant—the rich politicians are preferred by 1.47 points, a highly significant difference. Among poor White English speakers exposed to a speech by a White English candidate—that is, holding both the racial and linguistic treatment constant—the rich candidate is preferred by 1.69 points, a difference that is significant at the 0.05 level, even though there are only 19 subjects in this strata.<sup>37</sup> Among poor Blacks, rich politicians are preferred by 0.34 points, though this estimate is not significant; nor are rich Black Zulu or Xhosa candidates significantly preferred by poor Black Zulus or Xhosas, respectively, though again, sample sizes are small. In sum, the evidence suggests that poor voters favor rich candidates over poor candidates to a greater extent than even rich voters, and that this effect is particularly strong for poor Whites.

Yet, *why* do poor voters favor rich candidates? This is one important question for further discussion and interpretation. In principle, the additional outcome variables listed in Table 3—which stem from additional post-treatment questions asked after subjects had

---

<sup>37</sup> Among poor White Afrikaans subjects exposed to White Afrikaans candidates, the rich candidate is preferred by 0.96 points, though this difference is not statistically different from zero.

viewed the videotaped speeches—may give us some clues; for instance, they might allow us to assess whether poor subjects view rich candidates as more trustworthy, more likeable, more competent, and so on, or whether other variables distinguish rich and poor candidates. Figure 1 shows differences of mean evaluations of candidate attributes for poor subjects, that is, evaluations of mean perceived attributes of rich candidates minus evaluations of mean perceived attributes of poor candidates.

[FIGURE 1 ABOUT HERE]

Unfortunately, as Figure 1 makes clear, the answer for the moment is akin to “all of the above.” Indeed, poor subjects view rich candidates as more intelligent, competent, and impressive; as being more likely to keep their promises, do a good job in office, defend others, and successfully face the challenges of offices; and as having good motivations for running. Only on one variable, likeability, are rich candidates not judged more favorably than poor candidates (and there the difference is significant at the 0.1 level). This is in sharp contrast to the case of rich subjects shown in Figure 2, who for the most part do not view these attributes of rich candidates significantly favorably than they do poor candidates.<sup>38</sup>

[FIGURE 2 ABOUT HERE]

## **5. Conclusion**

Scholarship on South Africa has focused centrally on the role of race and, to a secondary extent, social class. Yet from electoral and survey data, it is not readily

---

<sup>38</sup> There are two differences that are statistically different from zero in Figure 2, that for likeability and for whether the politician would keep his promises. These would not survive a standard Bonferroni correction for multiple comparisons, however; that is, they could well be due to chance variation.

apparent how race and class interact with each other to shape voter preferences. Nor does the existing evidence tell us much about the role of language cleavages, which played an important political role during the transition to democracy and which may again help to shape South African politics.

This working paper takes a first pass at analyzing data from an experiment in which subjects recruited from a probability sample of residents of Johannesburg and Cape Town were shown similar videotaped political speeches in which the race, native language, and stated class background of the politician were varied. The results suggest that race does matter, but it matters much more centrally for South African Whites than for Blacks. However, there is some evidence that language matters as well, particularly for language groups such as Zulu- and Afrikaans-speakers who have arguably played a more marginal political role in the post-apartheid order than Xhosa- or English-speakers, respectively. Most surprisingly, the effects of class are not straightforward: for instance, poor subjects appear to favor rich candidates over poor candidates to a much greater extent than rich subjects, not just as a matter of vote intention but also when evaluating a host of candidate attributes.

The next task is to delve further into trying to explain these patterns. The experimental data itself should help here (for example, through analysis of additional outcome measures such as those in Table 3), and this analysis should be merged with analysis of electoral data as well as additional fieldwork. Analysis of data on the perceptions by Coloured respondents of the candidates might also help to shed light on this potentially pivotal demographic group.

## References

- Alexander, Neville. 2011. "After Apartheid: The Language Question." In Ian Shapiro and Kahreen Tebeau, eds., *After Apartheid: The Second Decade*. University of Virginia Press.
- Butler, Anthony. 2011. "Black Economic Empowerment since 1994: Diverse Hopes and Differentially Fulfilled Aspirations." In Ian Shapiro and Kahreen Tebeau, eds., *After Apartheid: The Second Decade*. University of Virginia Press.
- Carey, John M. and Matthew Soberg Shugart. 1995. "Incentives to Cultivate a Personal Vote: a Rank Ordering of Electoral Formulas." *Electoral Studies* 14 (4): 417-39.
- Dunning, Thad. 2009. "The Salience of Ethnic Categories: Field and Natural Experimental Evidence from Indian Village Councils." Working Paper, Department of Political Science, Yale University.
- Dunning, Thad, and Lauren Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104 (1): 21-39.
- Feree, Karen E. 2004. "The Micro-Foundations of Ethnic Voting: Evidence from South Africa." Afrobarometer Working Paper No. 40.
- Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945-960.
- Lieberman, Evan S. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. New York: Cambridge University Press.

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

Seekings, Jeremy. 2011. "Poverty and Inequality After Apartheid." In Ian Shapiro and Kahreen Tebeau, eds., *After Apartheid: The Second Decade*. University of Virginia Press.

Shapiro, Ian and Kahreen Tebeau, eds. 2011. *After Apartheid: The Second Decade*. University of Virginia Press.

Taylor, Michael, and Douglas Rae. 1969. "An Analysis of Crosscutting between Political Cleavages." *Comparative Politics* 1 (4): 534-547.

Wood, Elisabeth. 2000. *Forging Democracy from Below: Insurgent Transitions in South Africa and El Salvador*. New York: Cambridge University Press.



## Appendix: The role of candidate fixed effects

In this appendix, I discuss one methodological issue raised in the text—namely, the difficulty in manipulating race and language and thus in separating true race and language effects from race-independent or language-independent candidate fixed effects. Did the strategy of pooling across several actors help to elide this problem?

Table 9 presents, for all twelve actors, the mean evaluation (on the voting preference question) by White subjects (first column) and Black subjects (second column); the difference of mean evaluations by White and Black subjects, for each actor, appears in the third column.

Several observations can be made. First and perhaps least importantly, Black subjects rate all the candidates higher on average, across all twelve (White and Black) candidates, than do the White subjects: the average evaluation for White subjects is 3.46, while for Black subjects it is 4.33, for a highly significant difference of -0.87 ( $t$ -stat 8.05) (first row of Table 9). This is true for almost all of the individual actors as well, so that the entries of the final column are all negative (save one); however, many of the differences across White and Black subjects are not significant, due to the relatively small sample sizes for some of the actors. Secondly, however, White subjects evaluate Black candidates much more negatively than White candidates, while Black subjects tend to evaluate White candidates as positively as they do Black candidates. Indeed, on average, Black subjects rate white actors at 4.33 and also rate Black actors at 4.33, for a difference of zero (the numbers are exactly the same due only to rounding); yet, White subjects rate Black actors at 3.15 and White actors at 3.70, for a statistically-significant difference of -0.55 (which is about one-third of the standard deviation of evaluations by White

subjects). Thus, the difference-in-differences estimate—that is, the difference in White subjects' evaluations of White and Black candidates, minus the difference in Black subjects' evaluations of White and Black candidates (bottom-right cell of the table)—is  $0.55 - 0.00 = 0.55$  (s.e. 0.21), a difference which is statistically-significant at the 0.01 level. This is, of course, the same average finding reported in Table 5.

Finally, and most importantly for present purposes, inspection of Table 8 suggests fairly consistent effects across individual actors: White subjects evaluate most individual White candidates more positively and most individual Black candidates less positively, though there are a few exceptions (such as Xhosa-speaker 1, who was also rated especially positively by Black subjects, or Afrikaans-speaker 3). For Blacks, on the other hand, there is no systematic difference, and the single most positive average evaluation is of a White politician—English-speaker 1. While this exercise cannot completely rule out the identification problems discussed above—it could simply be that we happened to recruit six competent and attractive White actors and six less-competent and less-attractive Black actors—the fact that the effects hold up to within-actor analysis is encouraging. Notice, for example, that the size of the difference between evaluations by Whites and Blacks is much larger and typically statistically significant for the Black actors, while it is smaller and insignificant for the White actors. As the final column of the table shows, evaluations of the actors by Black subjects are roughly independent of race, while evaluations by White subjects are strongly dependent on race. The evidence thus suggests that the strong race and somewhat weaker language effects we find are not simply an artifact of race- or language-independent candidate fixed effects.

**Table 1: Experimental Design**  
**(Black African and White subjects)**

	Subject and candidate share same language	Subject and candidate have different language
Subject and candidate have same race	N=179 (90 same income, 77 different income*)	N=217 (99 same income, 106 different income*)
Subject and candidate have different race	N.A.**	N=331 (142 same income, 174 different income*)

\* “Same income” means that the subject were both “rich” or both “poor.” For politicians, “rich” corresponds to the version of the speech shown; for subjects, “rich” implies they were in the top 3 of 9 self-reported income categories. In each cell, the sum of subjects assigned to the “same income” and “different income” conditions does not total the number of subjects assigned to the corresponding race/language treatment, because of missing data on subjects’ income.

\*\* The experimental design assumes language is nested inside race, so this cell should be empty; see text and footnotes for discussion.

**Table 2: Racial, Linguistic, and Class Distribution of Survey Respondents**

Group	N	Percent of Total	Percent of Experimental Study Group
Poor White English-Speakers	64	3.4	10.1
Poor White Afrikaans-Speakers	28	1.5	4.4
Rich White English-Speakers	115	6.1	18.2
Rich White Afrikaans-Speakers	37	2.0	5.8
Poor Black Zulu-Speakers	155	8.2	24.5
Poor Black Xhosa-Speakers	134	7.1	21.2
Rich Black Zulu-Speakers	47	2.5	7.4
Rich Black Xhosa-Speakers	53	2.8	8.4
Black Speakers of Other First Languages*	312	16.5	--
Coloured English-Speakers	257	13.6	--
Coloured Afrikaans-Speakers	226	11.9	--
Indians/Asians	78	4.1	--
Other**	87	4.6	--
Did not report first language, income, or race	299	15.8	--
All	1,892	100	100

The table only includes respondents who reported their income category on a 9-point scale; respondents are coded as “rich” if they identify themselves in the top 3 of 9 income categories. First language and race are based on self-reports, using the Stats SA racial classifications (Black African, White, Coloured, Indian/Asian). Respondents in shaded cells are included in the main experimental analyses.

\*Among Blacks who do not speak Zulu or Xhosa as a first language (312 out of 701 Blacks), first languages include Afrikaans (35), English (43), Northern Sotho (55), Southern Sotho (75), Ndebele (13), Swati (24), Tsonga (16), Tswana (31), Venda (17) and other (3).

\*\* The “Other” category includes White respondents who speak a first language other than English or Afrikaans (22) and Coloured respondents who speak a first language other than English or Afrikaans (11).

**Table 3: Descriptive Statistics on Response Variables  
(Across All Treatment Conditions)**

<b>Variable</b>	<b>Range</b>	<b>Mean (SD)</b>	<b>Mean (SD) <i>Range 0-1</i></b>
Global Evaluation of Candidate	1-7	4.01 (1.87)	
Global Evaluation of Speech	1-7	4.38 (1.75)	
The Candidate...			
Is Likeable	1-5	3.44 (1.01)	
Is Intelligent	1-5	3.11 (1.02)	
Is Competent	1-5	3.10 (1.00)	
Is Impressive	1-7	4.15 (1.83)	
Would Keep His Promises	1-5	4.17 (1.68)	
Would Do a Good Job in Office	1-7	3.78 (1.54)	
Would Defend Others and Fight for His Ideals	1-7	4.31 (1.55)	
Has Good Motivations for Running	1-7	4.69 (1.59)	
Would Successfully Face Challenges of Office	1-7	4.16 (1.68)	
Has Good Ideas	1-7	4.94 (1.61)	

Descriptive statistics are presented for experimental subjects (Whites and Black Africans) only.

**Table 4. Race and Language Effects:  
Intention-To-Treat Analysis**

	Subject and candidate share same language	Subject and candidate have different language	Difference of means
Subject and candidate have same race	<b>A</b> 4.31 (0.14)	<b>B</b> 4.10 (0.12)	<b>A – B</b> 0.21 (0.18)
Subject and candidate have different race	<b>C</b> --	<b>D</b> 3.80 (0.11)	
Difference of means		<b>B – D</b> 0.30* (0.16)	<b>A – D</b> 0.51** (0.18)

Cells **A**, **B**, and **C** report average answers, by treatment assignment, to the question, “Please look at this ladder, which has seven steps. Suppose the first step of the ladder means ‘no, not at all’ and the 7th step means ‘yes, completely.’ Where would you put your answer to the following question: Does the speech make you want to vote for this candidate?”

Standard errors are in parentheses. \*  $p < 0.1$  \*\* $p < 0.01$

**Table 5. Race effects (sub-group analysis)**

	White Subjects	Black Subjects
Subject and Politician Have Same Race <b>A</b>	3.70 (0.11)	4.33 (0.09)
Subject and Politician Have Different Race <b>B</b>	3.15 (0.12)	4.33 (0.09)
Difference of Means <b>A – B</b>	0.55*** (0.16)	0.00 (0.13)

See notes to previous tables. \*\*\* p<0.001

**Table 6. Language Effects (sub-group analysis)**

	Afrikaans-speaking subjects	English-speaking subjects	Xhosa-speaking subjects	Zulu-speaking subjects
Subject and Politician Have Same Language <b>A</b>	4.07 (0.18) N=99	3.79 (0.13) N=166	4.54 (0.26) N=48	4.76 (0.24) N=55
Subject and Politician Have Different Languages <b>B</b>	3.70 (0.12) N=267	3.17 (0.09) N=463	4.46 (0.13) N=163	4.20 (0.15) N=162
Difference of Means <b>A – B</b>	0.37 (0.22)	0.63*** (0.16)	0.08 (0.27)	0.57* (0.30)

See notes to previous tables. This table combines the effects of language and race, because politicians who share the subject's language (and thus race) are compared to those who do not share the subject's language; the latter may have the same or a different race.

\* p<0.1 \*\*\* p<0.001

**Table 7. Within-Race Language Effects (sub-group analysis)**

	Afrikaans-speaking subjects	English-speaking subjects	Xhosa-speaking subjects	Zulu-speaking subjects
Subject and Politician Have Same Race and Language <b>A</b>	4.75 (0.42) N=20	3.55 (0.22) N=58	4.54 (0.19) N=48	4.76 (0.24) N=53
Subject and Politician Have Same Race, Different Language <b>B</b>	4.45 (0.37) N=20	3.44 (0.23) N=52	4.48 (0.26) N=62	4.30 (0.25) N=55
Difference of Means <b>A – B</b>	0.30 (0.55)	0.11 (0.32)	0.06 (0.32)	0.46 (0.35)

See notes to previous tables. Here, evaluations by subjects of a single race and language group of politicians who share their race and language group are compared to their evaluations of politicians who share their race but not their language group. This stratification results in quite small sample sizes.

**Table 8. The Effects of Social Class**

	Subject and Politician from the Same Income Group	Subject and Politician from Different Income Groups	Difference of Means
All subjects	3.77 (0.07)	3.99 (0.06)	-0.23* (0.10)
Poor subjects	3.88 (0.08)	4.21 (0.08)	-0.34** (0.11)
Poor White subjects	2.93 (0.25)	4.03 (0.24)	-1.10** (0.35)
Poor White English	2.25 (0.27)	3.46 (0.28)	-1.21** (0.39)
Poor White Afrikaans	4.00 (0.56)	4.64 (0.48)	-0.65 (0.74)
Poor Black subjects	4.40 (0.11)	4.61 (0.10)	-0.21 (0.15)
Poor Black Xhosa	4.84 (0.17)	4.38 (0.19)	0.46# (0.26)
Poor Black Zulu	4.24 (0.22)	4.67 (0.18)	-0.43 (0.29)
Rich subjects	3.48 (0.13)	3.42 (0.12)	0.07 (0.18)
Rich White subjects	3.62 (0.21)	3.09 (0.20)	0.53# (0.29)
Rich White English	3.60 (0.22)	3.12 (0.25)	0.48 (0.33)
Rich White Afrikaans	4.25 (0.52)	3.11 (0.36)	1.14# (0.61)
Rich Black subjects	3.80 (0.26)	3.93 (0.23)	-0.13 (0.35)
Rich Black Xhosa	4.4 (0.37)	3.36 (0.45)	1.04 (0.62)
Rich Black Zulu	4.0 (0.63)	4.2 (0.43)	-0.20 (0.73)

See notes to previous tables. Subject and politician are from the same income group if both are rich or both are poor; otherwise, they are from different income groups.

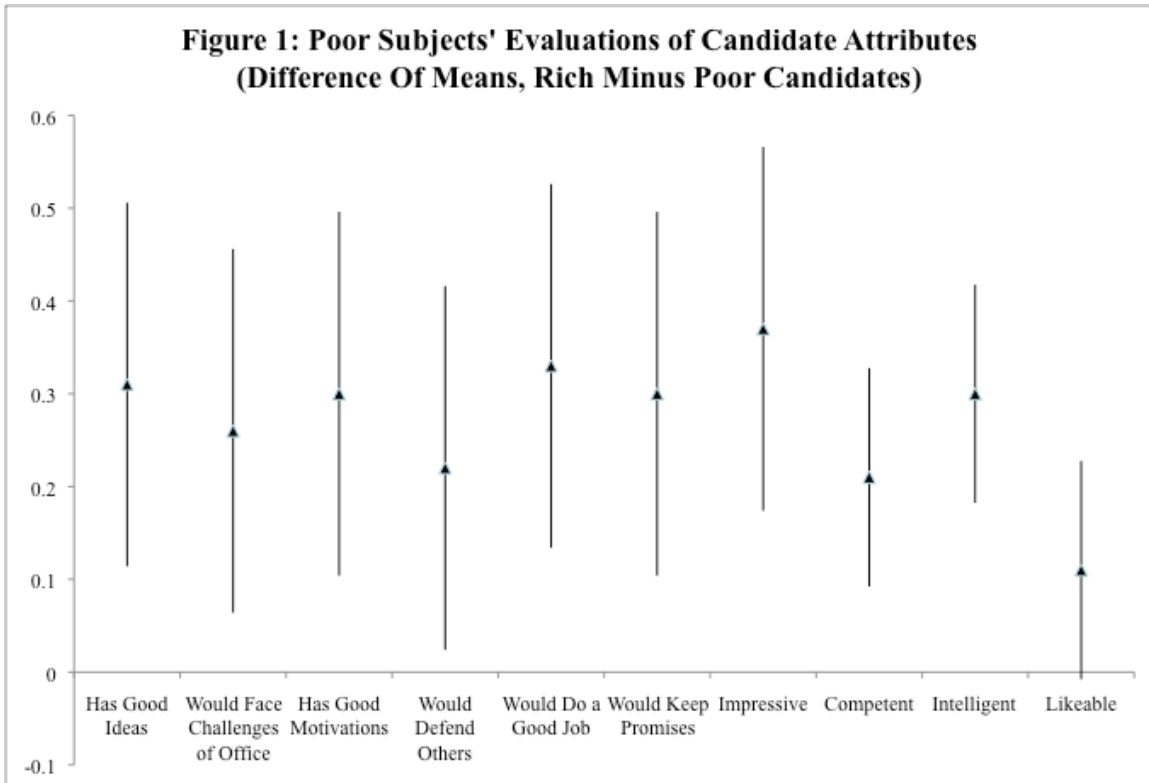
Standard errors are in parentheses. # p<0.1 \* p<0.5 \*\*p<0.01 \*\*\* p<0.001



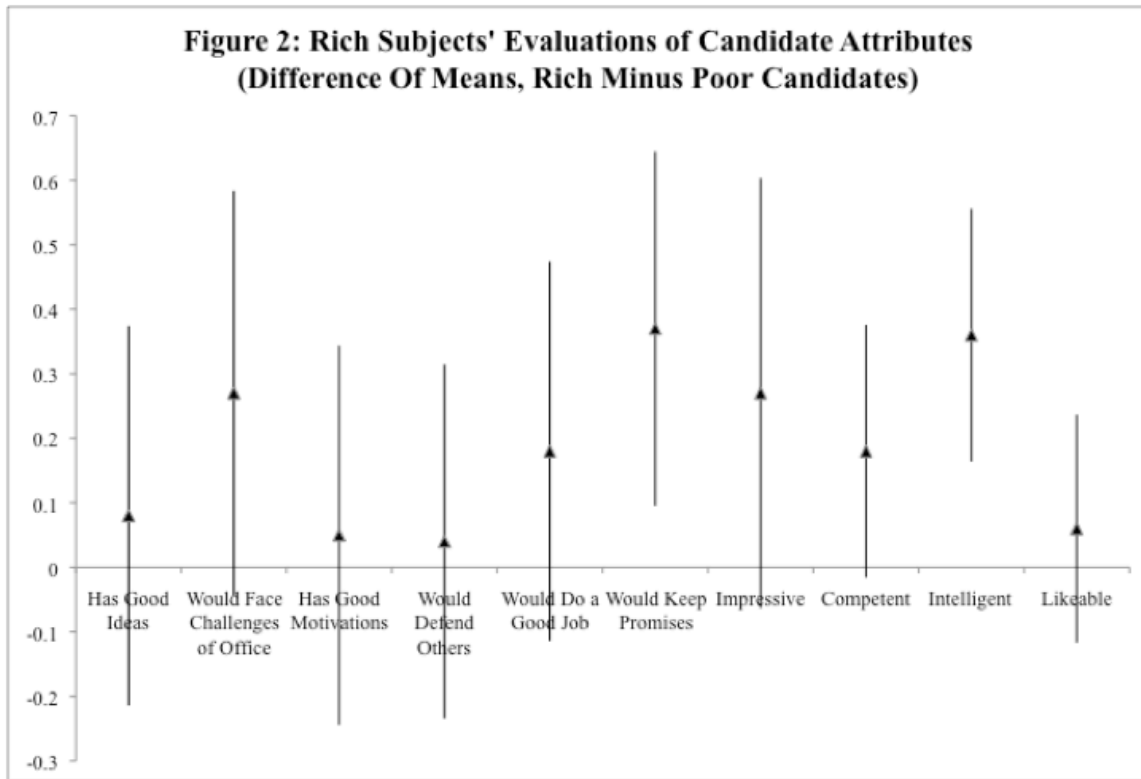
**Table 9. Race and language effects,  
by individual actors**

Race of Actor	Actor	White subjects (N) <b>A</b>	Black subjects (N) <b>B</b>	Difference of means (standard error) <b>A-B</b>
All actors		3.46 (411)	4.33 (773)	-0.87* (0.11)
White	English-speaker 1	4.5 (10)	5.13 (40)	-0.63 (0.60)
	English-speaker 2	3.57 (93)	4.31 (111)	-0.75* (0.23)
	English-speaker 3	4.0 (12)	4.60 (62)	-0.60 (0.58)
	Afrikaans-speaker 1	3.8 (25)	4.28 (88)	-0.48 (0.42)
	Afrikaans-speaker 2	3.84 (66)	3.81 (58)	0.04 (0.31)
	Afrikaans-speaker 3	3.21 (24)	3.94 (33)	-0.73 (0.56)
	All White actors (weighted average)	3.70 (230)	4.33 (392)	-0.63* (0.15)
Black	Zulu-speaker 1	3.21 (19)	4.58 (94)	-1.37* (0.43)
	Zulu-speaker 2	3.22 (76)	4.30 (106)	-1.08* (0.25)
	Xhosa-speaker 1	4.00 (3)	4.60 (42)	-0.60 (1.27)
	Xhosa-speaker 2	3.32 (49)	4.18 (55)	-0.86* (0.33)
	Xhosa-speaker 3	2.9 (10)	3.94 (48)	-1.04 (0.67)
	Xhosa-speaker 4	2.54 (24)	4.19 (36)	-1.65* (0.48)
	All Black actors (weighted average)	3.15 (181)	4.33 (381)	-1.17* (0.16)
Difference (White actors – Black actors)		0.55* (s.e. 0.16)	0.00 (s.e. 0.13)	0.55* (s.e. 0.21)

Cell entries give average answers to the question, “Please look at this ladder, which has 7 steps. Suppose the first step means ‘no, not at all’ and the 7<sup>th</sup> step means ‘yes, completely.’ Where would you put your answer to the following question: Does the speech make you want to vote for this candidate?” Only self-identified Whites and Black Africans are included. \*p<0.05



The figure reports the estimated effect of the candidate's income on evaluations of the candidate's attributes by poor subjects (those in the bottom 6 of 9 income categories). The darkened arrows give point estimates, while vertical lines show 95% confidence intervals. The analysis pools across race and language groups, for both subjects and politicians, so that mean evaluations by poor subjects of all poor candidates are subtracted from mean evaluations by poor subjects of all rich candidates.



The figure reports the estimated effect of the candidate's income on evaluations of the candidate's attributes by rich subjects (those in the top 3 of 9 income categories). The darkened arrows give point estimates, while vertical lines show 95% confidence intervals. The analysis pools across race and language groups, for both subjects and politicians, so that mean evaluations by rich subjects of all poor candidates are subtracted from mean evaluations by rich subjects of all rich candidates.