Appendix A

Contents:

1 Research Site Selection

The three townships, (Daveyton, Umlazi, and Mdantsane) were selected for inclusion in this study for three main reasons. First, these locations were selected as part of a larger study conducted in 2009 that needed representation across provinces and in local ethnic demography and thus the needs of the larger study also influenced our location selection. Therefore, while xenophobic violence was considered when selecting locations, it was not the only consideration, and we were limited to how influential it could be in our final selection decision (which is why we focus on municipal and provincial variation in experiences of xenophobic violence rather than township-level variation alone).

That said, each location does represent a different level of xenophobic attitudes as inferred by varying levels of violence throughout the wave of xenophobic attacks in May 2008. As Table A1 indicates, the three locations are representative of the variation we see in xenophobic attacks across **provinces** in May 2008. Daveyton represents an area within a municipality and province that were hard hit by xenophobic attacks, Umlazi an area within a municipality and province that experienced a moderate amount of attacks, and the area containing Mdantsane experienced a very lower number of attacks.

It is important to note that while we wanted variation in xenophobic violence across locations, we wanted to avoid the most violence locations. While asking questions about foreigners in a survey may not endanger our enumerators, the relatively unexpected start and quick spread of violence in 2008 (some of which could be opportunistic) motivated us to not select the most violent prone areas (e.g. Alexandria in Johannesburg or KwaMashu in eThekweni) out of concern for our research assistants and enumerators (many of which were female). In addition, not selecting the most violent and thus likely the most xenophobic areas we are likely estimating the determinants of relatively weaker xenophobia and so any effects found with our current sample are likely only more pronounced in more violent-prone areas.

The numbers in the Table A1 (below) were compiled from the International Organization for Migration (IOM) report (Misgao et al 2009), and the reader will note that these numbers are only estimates and actually group many attacks together. For example, if the report said "18 May: Attacks and looting continue in Jeppestown;

foreigners and minorities are told to leave. In Tembisa, 50 shacks are burned and four men are murdered. Seven arrests are made." (25), then we simply coded the 50 shacks being burned in Tembisa as one "attack" in Ekurhuleni Municipality and another attack in the Western Cape (Jeppestown). Because of the way the attacks were reported, we reported any number of attacks as one instance if they took place in the same location on the same day.

There is of course great variation across towns and neighborhoods within provinces and within municipalities, but we chose to look at these specific townships as representative of the areas that black South Africans tend to live in in each municipality (the modal Black South African in each municipality in the sampled areas lives in a township setting). These townships may not be perfectly representative of the municipalities they are a part of, but certain logistical matters dictated research in these areas. Daveyton was chosen because of its ethnically representative nature, its location in the Eastern part of Johannesburg where many attacks occurred, and because it was a safe place in which to conduct research on xenophobia (the ideal location would have been Alexandra where the attacks began and proliferated but that would not have been safe or feasible). Mdantsane was also convenient because the researchers had a number of contacts and experience conducting research there. Umlazi was also chosen because of its relatively safe nature in an area that suffered a number of attacks.

The variation and selection of the three townships is important because their variation in xenophobic attacks allows the research to understand xenophobia in the context of various levels of xenophobic attacks. One would expect to therefore see xenophobia to be highest in Daveyton and lowest in Mdantsane.

Third, the three research cites are representative of South Africa's townships. The three townships are all urban townships, with similar poverty and income levels, similar age and gender makeup, and similar average education levels (Statistics South Africa 2001).

The fourth selection criterion was the ethnic make-up of the townships. As Table 1 shows, Daveyton is ethnically diverse, while the other two cites are ethnically homogenous. This allows us to 1) isolate the impact of ethnic homogeneity on immigrant prejudice (i.e. one important aspect contributing to contact) and 2) study the two largest ethnic groups (Zulu and Xhosa) in South Africa in both homogenously and heterogeneously ethnic settings.

And finally, we were also limited in our ability to select townships by being able to find locations from which to conduct the larger 2009 study. Therefore, researcher connections and the receptiveness of location schools, universities, libraries, etc. to house the larger lab experiment in 2009 also influenced our location selection.

It is also important to note that we did not determine the race of respondents prior to participation. Given that the study was conducting in predominantly Black areas and we are interested in the attitudes of people living in these types of areas (and not only

black individuals), we did not feel it necessary to condition inclusion in the survey on race. However, as expected, very few (only four) respondents in the study did not self-identify as Black.

The map below will help the reader see where the three townships are in relation to each other. Daveyton is just outside Johannesburg, Umlazi just outside Durban, and Mdantsane is just outside East London. It takes about 8 hours to drive from Durban to either of the other locations, and it would take about 12 hours to drive from Johannesburg to East London.

Map A1

Table A1: Research Site Characteristics

Research Cite Townships	Location: Municipality, Province	Population of Municipality	Ethnic Makeup	Number of Xenophobic Attacks			
				In Township	In Municipality (outside of specific township)	In Province (outside of Municipality)	Total (Province)
Daveyton	Ekurhuleni (Eastern Johannesburg), Gauteng	2,724,227	Heterogeneous: Approximates national demography	1	21	35	57
Umlazi	eThekwini (Durban and surrounding area), KwaZulu-Natal	3,468,086	Homogenous: Zulu	1	9	2	12
Mdantsane	Buffalo City (East London and surrounding area), Eastern Cape	724,312	Homogenous: Xhosa	0	0	1*	1*
The Remaining 6 Provinces							Total
	Free State						2
	Limpopo						1
	Mpumalanga						1
	Northern Cape						0
	Northwest						3
	Western Cape						13

*This instance was rumored but never confirmed

2 Survey Environment 2009

The survey experiment presented in this paper was part of a larger field experiment concerned with ethnic identification. Each subject came to a central location in each of the three townships. They filled in the pre-study questionnaire, then saw 22 randomly selected videos of fellow South Africans and were asked to identify the ethnic identity of each individual. After the guessing was completed, each subject then filled out the post-study questionnaire, which includes the list experiment question analyzed in this study. Participants came to a central location and participated in the research at our "labs". In Daveyton and Umlazi we rented out a computer lab at a local community college, and in Mdantsane we rented out a computer lab at a primary school. These computer labs served as our research labs. The survey was largely conducted in English, but as needed, enumerators facilitated understanding by translating questions according to predetermined translations in Zulu, Xhosa, and Sotho. Rather than translate the survey into all eight major languages spoken in these areas, which in our experience reduces control over the similarity of questions across languages, we focused on ensuring the English was manageable for our study population and equipped our enumerators with the ability to consistently translate the questions in the three dominant languages as needed.

The research took place from July – August 2009.

Recruitment Procedures

Two local (preferably university students) will do the following:

a. The research team will identify randomly selected sections or areas within each township for the recruiters to canvass (door to door invitations):
    i. Selection process:
        1. Randomly select half the sections in each township.
        2. Within each section, randomly select three streets for canvassing.
        3. On each street select a home for a starting point on that street.
b. They will start at a designated street and home and follow an outlined path through the selected streets
c. They will knock on the door of every third house along that path and ask for the individual in the home who is 16 or older and had the most recent birthday. If this person is not home at that time, then they will leave the flyer with instruction for the person at the door to give the flyer to the person with the next birthday and have that individual contact us.
d. If the person is home, the recruiters will give them the flyer and the following:
    i. "A group of researchers from X University in the USA are here conducted a research project on ethnic identity. They want to find out what South Africans know about the cultures and ethnic groups in South Africa, and we would like you to participate."
    ii. We want to limit the information so that people are not primed before they go in to participate; we want everyone on the same page
    iii. They will then hand the person the flyer
    iv. To quell fears or help people see that it is easy and they do not have to be experts in culture, "The research consists of you viewing a few videos of different South Africans and guessing which tribe they come from. AS the

flyer says, you will receive the R20 for simply participating, and then you can win R30 from correctly guessing the tribes of the different people."

v. "Can we make an appointment for you to come to the [insert location] and participate?"
vi. They will then write the appointment time down on the back of the flyer and then record it on their schedule as well (with the person's real name, at least first name).
vii. The subject should bring the flyer to the research cite to prove that they were in fact recruited.

e. Should we also have the recruiters keep a tally of how many people they invite to participate so that we can see how large our non-response bias is? (this data was not collected)
f. It is your responsibility to make sure that your recruiters are clear on what they need to do and what they should and should not say.
g. Also, please let your recruiters know that we will do quality assurance checks to make sure they are recruiting in the way we want them to.

3 Survey Environment 2011

The differences between the 2009 and 2011 data gathering survey environments are 1) the venue 2) the experimental context and 3) the location of the survey, and 4) the survey questions. As with the 2009 survey, this survey was largely conducted in English, but as needed, enumerators facilitated understanding by translating questions according to predetermined translations in Zulu, Xhosa, and Sotho. Rather than translate the survey into all eight major languages spoken in these areas, which in our experience reduces control over the similarity of questions across languages, we focused on ensuring the English was manageable for our study population and equipped our enumerators with the ability to consistently translate the questions in the three dominant languages as needed.

First, in 2009, we recruited individuals to a central location because we needed to use computers for the indentifiability study. In 2011 while we recruited subjects in the same manner, we conducted the surveys in people's home. This resulted in a higher response rate and a more representative sample (see difference of means tests in the main paper). Conducted the survey at people's home may also provide people with a more comfortable environment in which to get more accurate responses. While the data gathered in 2009 is still reliable, it seems that the data from 2011 adds more reliability to the data set.

Second, the 2011 survey was not conducted in conjunction with the identifiability data. In 2011, it was only the survey that people completed. There was no computer image viewing or ethnic identifiability. This does introduce an interesting change to the data. It is possible that because people spent the majority of their time in 2009 guessing the ethnic identity of others, issues surrounding ethnic identity were primed and more salient. In addition, since the MEIM was administered before the list experiment in 2009, in 2011 subjects completed the MEIM at the end of the survey, to vary if their were any ethnic identity priming effects. Thus, if ethnic identity is primed in 2009, then we would see a significantly higher level of ethnic identity in 2009 relative to 2011, but we do not (see difference of means test in main paper).

The third major difference between the 2009 and 2011 data gathering process is the location of the survey. In 2009 three townships were surveyed, but in 2011 due to limitations in time and funds, only Daveyton was surveyed. Daveyton was chosen over the other two locations because it is more ethnically and socio-economically diverse relative to the other two townships. This allowed the second round of data gathering to include more members of the smaller ethnic groups and more people from more diverse socio-economic backgrounds in the most cost-effective manner.

Fourth, the survey questions in 2011 remained largely the same (the list experiment question remained exactly the same) as the survey questions in 2009. Some questions were deleted in the 2011 survey. This appendix contains the 2009 survey as it has the complete set of questions, but a copy of the 2011 survey is available upon request from the author.

4 Summary Statistics

Table A2: Covariate Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Positive Contact	1085	2.97	0.55	1	4
Daveyton	1088	0.69	0.46	0	1
Less than HS	1086	0.42	0.49	0	1
Completed HS	1086	0.38	0.49	0	1
Tertiary Education	1086	0.20	0.40	0	1
Private Water Source	1085	0.67	0.47	0	1
Low Income	1088	0.419	0.49	0	1
Unemployed	1088	0.782	041	0	1
Ethnic Strength	1085	3.26	0.45	1	4
Age	1080	31.48	13.45	16	79
Male	1087	0.50	0.50	0	1
Xhosa	1088	0.28	0.45	0	1
Zulu	1088	0.30	0.46	0	1
Other Ethnic	1088	0.42	0.49	0	1
Data Gathered in 2011	1088	0.45	0.50	0	1
Average Distance from Capitals	1088	523.9	134.3	449	828.5

Table A3: Proportion of Respondents per

	2009	2011
Daveyton	261	486
	(24%)	(45%)
Mdantsane	164	
	(15%)	
Umlazi	177	
	(16%)	

This table reports the sample size for each location-year of the survey, and the percentage of the overall sample from each location-year is in parenthesis.

5 Explanation of the Differences in the 2009 and 2011 Datasets

This section present the balance tests for the two rounds of data gathering and offers more detailed possible explanations for the differences in the datasets.

It is clear, from Table A4, that the two data sets are different. Subjects in the 2011 dataset, on average, have more positive contact, are more likely to have a private water source, are slightly older, and were angered by more elements in the list experiment when compared to those in the 2009 dataset. In addition, there are also more individuals with tertiary education in the 2011 dataset. Subjects in the 2009 dataset, on average, have stronger ethnic identities and are more likely to be Xhosa or Zulu than those in the 2011 dataset. However, in terms of gender composition and education level (except tertiary education), the two datasets are not significantly different (by conventional standards).

In the 2009 experiment, as mentioned above, all subjects answered the MEIM questions about ethnic identity after participating in another experiment that asked subjects to guess the ethnic identity of other South Africans; therefore, it is possible that ethnicity was primed in 2009. In 2011, the subjects did not participate in this identification experiment and thus ethnic identity may not be as strong because it was not primed. Also, the 2011 data has more diverse ethnic groups because Daveyton is more diverse than Mdantsane (Xhosa) and Umlazi (Zulu), and we surveyed 486 subjects from Daveyton in 2011 compared to 261 in 2009. Thus, it is expected that the 2011 subjects would be more ethnically diverse. It terms of age, it is possible that the 2011 subjects are slightly older than the 2009 subjects because in 2011 enumerators surveyed participants in their homes, but in 2009 we recruited subjects to a central location. It is possible that older participants, especially those over 50, would be more likely to participate if they could do so without going out of the house because of transport costs, health, etc. However, it is unclear exactly why subjects in 2011 would have more positive contact and higher education than those in the 2009 data.

The statistical difference with regards to the list experiment responses is a bit complex. Even when disaggregating the responses by treatment, the control group in 2011 is significantly greater than 2009, so it is not simply that people have become more xenophobic in the two years, although this is possible. It also seems that the control items, on average, have become issues that are angering more people since 2009. This is quite possible given that most of the list experiment control issues are economic in nature and that the worldwide recession still persists

and has worsened in South Africa in some respects between 2009 and 2011 (Mail and Gaurdian 2011; South Africa 2010).

Table A4: Difference of Means Tests
Testing for Balance on Observables Across 2009 and 2011 Data

Difference of Means Test: 2009 Data – 2011 Data								
			Means				Observations	
Variable	Min	Max	2009	2011	Difference	P-value	2009	2011
Positive Contact	1	4	2.90	3.06	-0.162	0.00	599	486
Less than HS	0	1	0.40	0.44	-0.041	0.18	601	485
Completed HS	0	1	0.36	0.40	-0.040	0.18	601	485
Tertiary Education	0	1	0.23	0.15	-0.080	0.00	601	485
Private Water Source	0	1	0.62	0.74	-0.126	0.00	600	485
Low Income	0	1	0.73	0.03	0.702	0.00	602	486
Unemployed	0	1	0.86	0.689	0.168	0.00	602	486
Ethnic Strength	1	4	3.29	3.21	0.083	0.00	599	486
Age	16	79	28.03	35.70	-7.666	0.00	594	486
Male	0	1	0.52	0.48	0.501	0.29	601	486
Xhosa	0	1	0.35	0.19	0.154	0.00	602	486
Zulu	0	1	0.37	0.21	0.164	0.00	602	486
Other Ethnic	0	1	0.28	0.60	-0.318	0.00	602	486
Anger Experiment Response	0	1	2.91	3.69	-0.784	0.00	602	486
Vote Experiment Response	0	1	3.10	2.89	-0.204	0.02	599	479

Table 1 reports the results of difference of means tests on observable covariates between the 2009 and 2011 data sets. The table shows that there are significant differences on most observables. The two data sets are not significantly different in terms of education (Less than HS, Completed HS) and gender (male).

6 Robustness to Pooling Alternatives

Given that the two datasets differ in a number of ways, it is important to ensure that the pooling of these datasets does not drive the results and that pooling the data is reasonable. We therefore run two sets of t-tests. We first test to see if the main results regarding unemployment (economic threat) hold for the 2009 and the 2011 samples individually. We then test to see if the results hold for the Daveyton sample, which is the only location for which we have two years of data and for the pooled Mdantsane/Umlazi sample (locations with only 2009 data). If the results holds then we can be sure that 1) the patterns we observe are consistent across years and 2) that the patterns hold if we only consider the location for which we have multiple year data, and 3) that the results are not driven by Daveyton alone (which accounts for 69% of the sample). We only run t-tests because power is not sufficient to reliably run the ICT analysis on each sub-sample; in fact the motivation to gather more data in 2011 was driven by the need to increase power in order to run ICT estimation.

The results, and the main results from Table 2 (as a reference), are reported in Table A5 and confirm the main conclusions that unemployment (economic threat) drives anger over immigrants moving into one's neighborhood. In every case, there is a significant difference in response rates between the control and treatment groups for the unemployed but never for the employed. These results are likely only driven by the small sample size of employed respondents in the Mdantsane/Umlazi analysis (given the relatively high standard errors associated with these estimates). The results for the vote experiment hold in three of the four sub-samples but given that this result is not robust to ICT estimation, we are not worried about the weaker and less consistent results from the vote experiment.

In addition to this analysis, we also estimated Chow tests in order to determine if there is fact a structural break between the 2009 and 2011 data. In other words, the Chow test indicates if the data can reasonably be pooled into a single regression or if the regression line resulting from a pooled analysis would overlook changes in the regression line due to differences in the two samples. We used the STATA package "chowreg" to estimate chow tests on the regression models estimated using ICT. We ran chow tests for four different regressions: the treatment and control group separately for each list experiment. We ran separate regression because we are particularly concerned that we in fact might have two different treatment groups and two different control groups that cannot be reasonably pooled together. The results of the test reveal that there is no evidence of a structural break in any instance and thus we can reasonably pool the data for 2009 and 2011. More specifically, we fail to reject the null hypothesis of no structural breaks because the p-value for the F statistic for each regression model is greater than the conventional .05.

Table A5: Analysis using Alternative Pooling

	Anger Experiment		Vote Experiment		Difference:	
	Control	Treatment	Control	Treatment	Anger Toward Immigrants	Voting Against Immigrants
	(1)	(2)	(3)	(4)	(5: 1-2)	(6: 3-4)
Main Results						
Unemployed	2.82	3.58	2.83	3.10	-0.77***	-0.26**
	(0.07)	(0.08)	(0.08)	(0.09)		
	271	286	286	285		
Employed	3.28	3.45	2.69	3.04	-0.18	-0.34
	(0.11)	(0.16)	(0.15)	(0.17)		
	85	83	72	82		
2009						
Unemployed	2.45	3.21	2.97	3.23	-0.76***	-0.27*
	(0.09)	(0.11)	(0.10)	(0.12)		
	161	179	172	179		
Employed	2.9	3.03	2.81	2.97	-0.13	-0.15
	(0.21)	(0.36)	(0.27)	(0.32)		
	30	29	27	31		
2011						
Unemployed	3.35	4.21	2.64	2.87	-0.85***	-0.23
	(0.09)	(0.11)	(0.14)	(0.15)		
	110	107	114	106		
Employed	3.49	3.68	2.62	3.08	-0.19	-0.46
	(0.12)	(0.19)	(0.19)	(0.21)		
	55	54	45	51		
Daveyton						
Unemployed	3.05	3.86	2.82	3.08	-0.81***	-0.26*
	(0.08)	(0.10)	(0.10)	(0.11)		
	182	176	184	182		
Employed	3.39	3.52	2.74	3.00	-0.12	-0.25
	(0.11)	(0.17)	(0.16)	(0.19)		
	71	69	59	67		
Mdantsane/Umlazi						
Unemployed	2.34	3.15	2.87	3.14	-0.81***	-0.26**
	(0.13)	(0.14)	(0.13)	(0.16)		
	89	110	102	103		
Employed	2.71	3.14	2.46	3.20	-0.43	-0.73
	(0.32)	(0.42)	(0.43)	(0.45)		
	14	14	13	15		

7 Balance on Observables Across Treatment and Control

Before running these tests on the list experiment data, we also check that the sampling procedures achieved a balance on pre-treatment observables, which ensures that random assignment to treatment is not correlated with any of the covariates of interest and is thus exogenous to the model. Tables A6 and A7 show results from difference of means tests, which test for a significant difference between the control group mean and each treatment group mean for the list experiment. Tables A6 and A7 show that random assignment did in fact balance on the majority of pre-treatment observables for both the Anger Experiment (except for positive contact and Zulu ethnicity) and Vote Experiment. It is important to interpret the results of these tests with the caveat that those in the treatment conditions have slightly higher average levels of positive contact than those in the control group.

Table A6: Random Assignment to Treatment
Balance on Observables: Anger List Experiment

Panel A: Immigrants Moving In Treatment

			Means				Observations	
Variable	Min	Max	Control	Immigrant	Difference	P-value	Control	Immigrant
Positive Contact	1	4	2.92	3	-0.076	0.07	355	369
Daveyton	0	1	0.71	0.66	0.047	0.18	356	369
Less than HS	0	1	0.39	0.43	-0.040	0.27	355	368
Completed HS	0	1	0.42	0.38	0.042	0.25	355	368
Tertiary Education	0	1	0.19	0.19	-0.001	0.96	355	368
Private Water Source	0	1	0.67	0.67	-0.002	0.94	356	368
Low Income	0	1	0.42	0.43	-0.004	0.92	356	369
Unemployed	0	1	0.76	0.78	-0.014	0.66	356	369
Ethnic Strength	1	4	3.27	3.25	0.021	0.53	355	369
Age	16	79	31.2	31.1	0.147	0.88	355	366
Male	0	1	0.53	0.49	0.034	0.37	356	368
Xhosa	0	1	0.28	0.27	0.013	0.70	356	369
Zulu	0	1	0.28	0.34	-0.058	0.09	356	369
Other Ethnic	0	1	0.44	0.40	0.045	0.22	356	369
Data From 2011	0	1	0.46	0.44	0.027	0.46	356	369

Table A7: Random Assignment to Treatment
Balance on Observables: Vote List Experiment

Panel B: Anti-Immigrant Candidate Treatment								
			Means				Observations	
Variable	Min	Max	Control	Candidate	Difference	P-value	Control	Candidate
Positive Contact	1	4	2.96	2.98	-0.016	0.69	359	370
Daveyton	0	1	0.68	0.68	-0.003	0.93	359	372
Less than HS	0	1	0.41	0.41	0.002	0.96	358	372
Completed HS	0	1	0.39	0.40	-0.012	0.74	358	372
Tertiary Education	0	1	0.19	0.19	0.010	0.73	358	372
Private Water Source	0	1	0.68	0.65	0.037	0.28	358	371
Low Income	0	1	0.41	0.44	-0.034	0.35	359	372
Unemployed	0	1	0.80	0.77	0.025	0.41	359	372
Ethnic Strength	1	4	3.27	3.25	0.016	0.62	359	370
Age	16	79	32.2	30.8	1.474	0.15	358	367
Male	0	1	0.48	0.51	-0.036	0.33	358	372
Xhosa	0	1	0.26	0.28	-0.020	0.54	359	372
Zulu	0	1	0.31	0.33	-0.013	0.70	359	372
Other Ethnic	0	1	0.43	0.39	0.034	0.35	359	372
Data From 2011	0	1	0.45	0.44	0.010	0.78	359	372

Appendix B

MEIM (Multi-ethnic Identity Measure) Questions
Calculation procedures for the contact and ethnic strength scores

Ethnic/Culture Identity Questions (MEIM Questions; adapted from Phinney 1992)

In South Africa, people come from a lot of different cultures and there are many different words to describe the different backgrounds or ethnic groups that people come from. Some examples of the names of ethnic groups are Xhosa, Zulu, Tswana, Afrikaans, English, and Coloured. Every person is born into an ethnic group, or sometimes two groups, but people differ on how important their ethnicity is to them, how they feel about it, and how they act because of it. These next questions are about your ethnic/culture group and how you feel about it or react to it.

Please fill in:
In terms of ethnic group, I am (circle one: Ndebele, Pedi, Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa, or Zulu).
Use the numbers given below to show how you agree or disagree with each statement.
4: Agree Very Much with the statement
3: Agree A Little with the statement
2: Disagree A Little with the statement
1: Disagree Very Much with the statement

1. I have spent time trying to find out more about my own ethnic group, like its history, traditions, and customs _____
2. I am active in organizations or groups that include mostly members of my own ethnic group _____
3. I have a clear sense of my ethnic background and what it means for me _____
4. I like meeting and getting to know people from ethnic groups other than my own _____
5. I think a lot about how my life will be affected by my ethnic group membership _____
6. I am happy that I am a member of the group I belong to _____
7. I sometimes feel it would be better if different ethnic groups didn't try to mix together _____
8. I am not very clear about how my ethnicity affects my life _____
9. I often spend time with people from ethnic groups other than my own _____
10. I really have not spent much time trying to learn more about the culture and history of my ethnic group _____
11. I have a strong sense of belonging to my ethnic group _____
12. I understand pretty well what my ethnic group membership means to me, in terms of how to relate to my own group and other groups _____
13. In order to learn more about my ethnic background, I have often talked to other people about my ethnic group _____
14. I have a lot of pride in my ethnic group and its accomplishments _____
15. I do not try to become friends with people from other ethnic groups _____
16. I participate in cultural practices of my own group, such as special food, music, or customs _____
17. I am involved in activities with people from other ethnic groups _____
18. I feel a strong attachment toward my own ethnic group _____

19. I enjoy being around people from ethnic groups other than my own _____
20. I feel good about my cultural or ethnic background _____
21. My father's ethnic group is (circle one: Ndebele, Pedi, Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa, or Zulu).
22. My mother's ethnic group is (circle one: Ndebele, Pedi, Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa, or Zulu).

Score Calculation Procedures Using Subject Responses to the MEIM Questions

Contact score: first reversing responses to questions 2 and 15, then averaging the responses to 2r, 9, 15r, 17, and 19. This provides a score that is increasing in positive contact with other ethnic groups.

Strength of ethnic identity score: first reversing responses to questions 8 and 10 and then averaging across the response to 1, 3, 5, 6, 8r, 10r, 11, 12, 13, 14, 16, 18, and 20. This creates a measure that is increasing in the strength of ethnic identity.

Appendix C

Testing the "No Design Effects" Assumption of Anger List Experiment and a Discussion of Addressing the "No Liars" Assumption

In order to test the no design effects assumption, we first estimate the proportion of individuals who answered affirmatively to the sensitive item in each experiment (see Blaire and Imai 2011). Table C1 estimates the proportion of individuals who responded affirmatively to the sensitive item conditional on the individual's response to the list experiment (y). The first column lists the observed y-value, the columns marked z = 0 represent a negative response to the sensitive item, and columns marked z = 1 represent an affirmative response to the sensitive item. Standard errors of each estimate are also reported[1].

Table C1, for example, shows that for the immigrant treatment in the anger experiment an estimated 11.7% of subjects who have a dependent variable of y = 3 responded negatively to the sensitive item and that 14.8% of respondents with a dependent variable of y = 3 responded affirmatively to the sensitive item. Table C1 reports that an estimated 62.8% of individuals in the immigrant treatment group responded affirmatively to the sensitive item and that 21% of subjects in the candidate treatment of the vote experiment responded affirmatively to the sensitive item.

The reader will note that some of the estimates in Table C1 are negative. Blaire and Imai (2011) propose a test to determine the likelihood of observing a negative estimate in the presence of the assumption of no design effects. The assumption of no design effects states that the addition of the sensitive item does not change a subject's response to the control items.

For example, in the anger list experiment, it could be argued that "The unemployment rate" control item would be correlated with people's responses to their anger over immigrants (given the anecdotal evidence that people are angered by immigrants because they steal jobs). Due to this correlation, the presence of the "Immigrants moving into your neighborhood" sensitive item would alter people's response to the unemployment control item (for example if one saw the control list, unemployment may not anger them until they see it in conjunction with the immigrant item). A violation of the no design effects assumption, such as the previous example, would make it difficult to accurately compare the control and treatment groups do to heterogeneous treatment effects. Therefore, we test for the existence of design effects using Blaire and Imai's (2011) proposed test. The results yield a Bonferroni-corrected p-value for the immigrant treatment of 0.849 and for the candidate treatment a p-value of 0.582. These values are both well above the threshold of $\propto = 0.05$ and thus we fail to reject the null hypothesis of no design effects, and we can be relatively confident that both experimental treatments are free of design effects and the first assumption holds (see Blair and Imai 2011 for further details on the limitations of this test).

We address the "no liars" assumption by controlling for ceiling and floor effects in the ICT estimation. The "no liars" assumption states that people will honestly respond to the sensitive item. This assumption can be addressed in the proposed maximum likelihood estimation by controlling for ceiling and floor effects. Ceiling effects occur when subjects agree with, or answer affirmatively to, all control items (J = number of control items) and the sensitive

[1] Standard errors are not reported for cases where y = 0 and z = 0 (meaning the subject responded to the list question with 0 and they answered negatively to the sensitive item) because this case is directly observed and does not need to be estimated.

item, but subjects respond to the list question with y = J rather than y = J + 1 (Blaire and Imai 2011). Floor effects occur when a respondent's true view is only responding in the affirmative to the sensitive item and not to any of the control items, but they respond with y = 0 because they don't want to signal that they are only angered by the sensitive item. This is most probable when the control items are more likely to generate many negative responses (Blaire Imai 2011). In both instances, subjects feel that their response to the sensitive item will be detected when answering with their true preferences. Therefore, subjects underreport their true feelings by excluding the sensitive item from their response.

In these list experiments, ceiling effects are more of a concern. In the list experiments, few respondents in the treatment conditions reported 0 in response to the list question (2.2% in the anger experiment and 5.2% in the vote experiment). Also, the control items in the list are issues that are relevant to individuals in South Africa's townships and are likely to generate many affirmative responses rather than many negative responses. Thus, ceiling effects are an important concern, which we address in the estimation. Therefore, we estimate maximum likelihood models using Blaire and Imai's (2011) ICT method controlled for ceiling effects.

Table C1: Estimated Population Proportions for Anger and Vote List Experiment

	Anger: Immigrants				Vote: Candidate			
y value	z = 0	s.e.	z = 1	s.e.	z = 0	s.e.	z = 1	s.e.
0	2.2%	N/A	-0.2%	1.1	2.2%	N/A	2.0%	1.3
1	10.0%	1.8	4.9%	2.6	9.0%	2.1	3.0%	2.6
2	5.9%	3.0	8.1%	3.3	8.8%	3.0	4.8%	3.3
3	11.7%	3.6	14.8%	3.7	21.5%	3.5	-2.0%	3.7
4	7.5%	3.6	35.2%	2.5	35.6%	3.2	-9.2%	2.9
Total	37.3%		62.8%		77.1%		21.0%	

Table C1 estimates the proportion of individuals who responded affirmatively to the sensitive item given the number of control items they also answered affirmatively. The first column lists the observed y-value (the subject's response to the list experiment question), columns marked z = 0 represent a negative response to the sensitive item, and columns marked z = 1 represent an affirmative response to the sensitive item. Standard errors of each estimate are also reported (standard errors are not reported for cases where y and z are both 0 -- meaning the subject responded to the list question with 0 and they answered negatively to the sensitive item -- because this case is directly observed and does not need to be estimated).