

Supplementary Materials

Who Participates in Communal Riots? Survey Evidence from South Africa

Contents

1	Conducting the Survey	A-2
1.1	Initial Preparations	A-2
1.2	Translation	A-2
1.3	Sampling	A-2
2	Question Wording and Measurement	A-6
3	Missing Data	A-8
4	Additional Results	A-9

List of Tables

A-1	Comparison of the Sample With Previous Population Estimates	A-4
A-2	Descriptive Statistics	A-6
A-3	Item Wording, Missing Values and Scale Construction	A-7
A-4	Participation in 2008 by Riot Behaviour and Gender	A-9
A-5	Ordered Logit Regressions of Participation in 2008	A-10
A-6	Logit Regressions of Participation in 2008 With Missing Values Dropped	A-11

List of Figures

1. Conducting the Survey

1.1. Initial Preparations

The fieldwork commenced with exploratory, open-ended interviews, conducted with Alexandra residents in August 2010. The research site was defined as “Old Alexandra”: the traditional area of Alexandra as demarcated by Wynberg Road in the south, Vasco Da Gama Street in the north, 1st Street in the west, and the Jukskei river in the east (Bonner and Nieftagodien 2008). In other words, I excluded the newer neighbourhoods of government housing to the east of the Jukskei river. The survey fieldwork then took place in April and May 2011. Interviewers were recruited using a contact at a Johannesburg survey firm. Four interviewers were hired; all were all black South Africans who lived in Johannesburg (none resided in Alexandra); three were male and one female; and each had at least three years’ experience.

1.2. Translation

Alexandra is a linguistically diverse township. Respondents reported ten different first languages. Fortunately, the African languages spoken in South Africa fall into four internally homogenous groups, where the speakers of a language within one group are typically able to understand speakers of another language within the same group (Lewis 2009). Moreover, two of the language groups—the Nguni group featuring Zulu, Xhosa, Ndebele, and Swazi, and the Sotho group featuring Sotho, Pedi, and Tswana—are very large.¹

The survey instrument was thus translated into two languages: one from the Nguni group (Zulu) and one from the Sotho group (Tswana). Existing survey data² show that Zulu is the most popular language in Alexandra. More importantly, it is the *lingua franca* in Johannesburg townships. Tswana was chosen because it was the language from the Sotho group that the interviewers were most comfortable with.³

I began the process of translation during the pre-fieldwork training period. I used the method of decentering (Werner and Campbell 1970), where changes are made between the original and target languages until all versions are satisfactory. The questions were read aloud in Zulu and Tswana. Translations were recorded. Ambiguities and disagreements were discussed, and if necessary, the original English wording was adjusted. Working off the notes made during this training period, the instrument was translated into Zulu and Tswana by teams of two interviewers chosen according to their fluency with the target language.

1.3. Sampling

Drawing a representative sample in a South African township such as Alexandra is challenging. There are no reliable lists of residents, households or dwellings, few people have fixed-line telephones and even fewer have access to the internet. This challenge is

¹A large majority of Alexandra residents (89%) reported speaking a language from one of these groups as a first language.

²See Table A-1 for further details.

³Two of the interviewers were native Tswana speakers and thus adept with Sotho and Pedi. They also had a good command of Zulu. The other two were Zulu-speakers who were comfortable with using the Tswana instrument.

overcome by stratifying the sample by housing type. The reasoning is that, although population density varies across housing types, it does not vary much within housing types. Satellite photographs could thus be used to select clusters within each housing type. Existing estimates of the Alexandra population within each housing type, supplied by the Alexandra Benchmark Survey of 2005,⁴ are then used to adjust for the differing probabilities of selection across housing type strata.

Three housing types are used as strata. The first is the fairly regular mixture of small single-storey houses, shacks, and backyard rooms that constitutes the vast majority of Alexandra's housing stock. There are also two other smaller, but distinct housing types: government flats, and the hostels. The sample was also stratified by gender at the point of respondent selection to guard against undersampling men. Population data for gender were also obtained from the Alexandra Benchmark Survey.

Starting points, or more formally, clusters, were selected by dividing each housing type into equal-sized chunks using a satellite photograph of Alexandra—a task that was made substantially easier by the grid street pattern of the area. There were two hostel clusters, three clusters of flats and 80 house-and-shack clusters. The Alexandra Benchmark Survey estimated that 92% of residents live in the area of houses and shacks. Eighteen out of the 20 clusters were therefore selected from this area, with only one each from the hostels and apartments. Selection was done by numbering the clusters and using a random number generator.

This stratified, clustered design permits the random selection of dwelling units—individual houses, shacks, apartments, and hostel rooms. Households and respondents are then selected within these dwelling units.⁵ The following method was used. Every day in the field corresponded with a particular cluster. Each interviewer attempted to contact nine respondents per day / cluster, producing a targeted sample of 720. A typical procedure was for each cluster to be divided in four, with one interviewer assigned to each segment. Interviewers first sampled dwelling units (house, shack, flat or hostel room) by walking in a given direction and using a sampling interval that varied between three and seven. Upon making contact with someone in a selected dwelling, interviewers asked for the number of households sharing that house (and its yard), shack, flat or hostel room. The household was selected using the same sampling interval used to choose dwellings. Within each household, interviewers selected a South African⁶ aged eighteen years old as a respondent. A random number table was used.⁷ The sample was also stratified by gender at this stage, with interviewers alternatively seeking men and women to interview. The survey instrument was administered by interviewers using a face-to-face method.

⁴The Alexandra Benchmark Survey (ABS) is a large survey of Alexandra households, commissioned by The Alexandra Renewal Project, an agency in the Gauteng Province Department of Housing. It was designed to provide estimates of development and quality of life indicators for policymakers (Alexandra Renewal Project ND). The ABS is preferred to the 2001 Census as a source of population estimates for two reasons. First, their estimates are more recent than the census'. Second, the ABS provides demographic marginals for the area of Old Alexandra—my survey site—while census data is only available for a slightly larger area that includes some newer blocks of government housing.

⁵A household was defined for respondents as all the individuals who regularly eat from the same pot. This technique is used in the Afrobarometer survey project (Afrobarometer Network 2007). One house was found to be occupied by 20 households, both inside, and outside in backyard shacks and rooms.

⁶If the household consisted only of foreign nationals, the interviewer apologized and counted this household as not eligible for the sample.

⁷This table featured rows for the date of interview and columns for the number of people of the appropriate gender living in the household. The number, n , in the relevant cell of the table gave the n -th oldest man/woman to be selected as a respondent.

Respondents' answers were recorded by interviewers on paper questionnaires.

Table A-1. Comparison of the Sample With Previous Population Estimates

	ABS 2005 Survey ^a	2011 Survey			
		2008 Subsample ^b		Full Sample	
	(%)	(%)	(S.E.)	(%)	(S.E.)
Gender					
Male	52.8	52.6	(1.7)	52.9	(1.5)
Female	47.2	47.3	(1.7)	47.1	(1.5)
Employment Status					
Employed	50.0	43.8	(3.2)	43.7	(3.2)
Not in the labour force	19.5	10.9	(1.4)	10.8	(1.4)
Unemployed	30.5	45.3	(2.7)	45.5	(2.6)
Age					
18–24	20.0	16.4	(1.7)	18.4	(1.6)
25–34	32.3	33.8	(2.0)	33.7	(1.9)
35–49	27.7	33.3	(2.6)	32.1	(2.3)
50+	20.0	16.5	(1.9)	15.8	(2.0)
Language					
Zulu	32.0	36.4	(4.7)	36.3	(5.0)
Pedi	25.0	20.0	(2.9)	20.4	(2.9)
Tswana	13.0	10.8	(1.7)	10.9	(1.7)
Sotho	9.0	10.8	(1.6)	10.5	(1.5)
Xhosa	7.0	10.2	(2.1)	10.2	(1.8)
Other	14.0	11.7	(1.9)	11.6	(1.8)
<i>N</i>	1255 ^c	452		497	

^a 2005 Alexandra Benchmark Survey (Alexandra Renewal Project ND), subsample from Old Alexandra.

^b The subsample from my 2011 survey who reported living in Alexandra at the time of the violence in May 2008.

^c This is the number of households sampled. The Alexandra Benchmark Survey recorded data from all the members of a sampled household.

The estimates from the 2011 survey are weighted by the inverse of the probability of selection, but are not post-stratified. Standard errors (in parentheses) are design-based.

One call-back was required for non-contact at the level of dwelling units, household, or respondent, and no substitutions were permitted. Interviews were completed for 497 respondents. There were 83 instances of non-cooperation at the level of dwelling unit and household and a further 23 at the respondent level, 90 cases of non-contact at the level of dwelling or household and another four at the respondent level, 16 cases where non-contact occurred for other reasons, and six cases of non-contact due to safety concerns in one cluster.⁸ The response rate (AAPOR #3) for this survey is thus 69.2%.⁹

⁸A group of men began aggressively questioning the intentions of one interviewer in the “Beirut” area in policing sector 2, previously the site of partisan violence between the Inkatha Freedom Party (IFP) and the African National Congress (ANC). The team left the area immediately.

⁹AAPOR #3 is the response rate definition recommended by the Council of American Survey Researchers (CASRO) (Weisberg 2005). The formula is as follows:

$$RR3 = \frac{I}{(I + P) + (R + NC + O) + e(U)}$$

where I = complete interviews, P = partial interviews, R = refusals, NC = non-contacts, O = other eligible nonresponse, e = estimate of proportion of U that is eligible, and U = unknown eligibility.

This is a similar response rate to high-quality face-to-face surveys conducted these days in the United States such as the American National Election Study and the General Social Survey (Dixon and Tucker 2010). The achievement of such a response rate despite the small number of callbacks and the lack of respondent incentives can be attributed to two factors. First, a marked willingness to be interviewed among residents of the area.¹⁰ Second, 14 of the 21 days of fieldwork took place on public holidays or weekends. This reduced the level of non-contact due to people being at work.

Table A-1 compares the marginals of my survey to those from the Alexandra Benchmark Survey (ABS).¹¹ A comparison of the two sets of marginals permits a check of the adequacy of my sampling frame and sampling procedure.¹² The 2011 response distributions for age and language group correspond quite closely to those from the ABS, particularly my full sample. The marginals for employment status are not as concordant. My estimates of the proportion of people who are unemployed and not in the labour force are larger than those obtained by the ABS.

In addition to weighting my analyses by household size, gender and housing type, I also post-stratify using these ABS estimates of the proportion of Alexandra residents in each age, employment and language group.

¹⁰This willingness to participate in the survey produced an unforeseen problem: one interviewer reported that several individuals who were not selected into the sample nevertheless demanded to be interviewed. He conducted brief interviews to assuage their curiosity, and then moved on.

¹¹The latter results were not released with uncertainty estimates. Formal tests of difference of proportions are thus not possible. These results are useful however, because, firstly they are more up-to-date than the most recently available census data, which was gathered in 2001; secondly, the ABS results offer survey marginals for gender, housing type, employment status, age, and language within the area of Old Alexandra—the same population I sampled.

¹²Both sets of results are weighted. The results from my survey are weighted to adjust for the differential probabilities of respondent selection due to household size and the stratification by housing type and gender. No post-stratification is used.

2. Question Wording and Measurement

Table A-2. Descriptive Statistics

Variable Name	Mean	Std. Dev.	Min.	Max.
Participation in 2008	.09	.27	0	1
Party support: None	.33	.47	0	1
Party support: Opposition	.12	.33	0	1
Meeting attendance	.59	.98	0	3
Gender: Male	.52	.50	0	1
Education: Primary school	.13	.33	0	1
Education: Some high school	.37	.48	0	1
Education: Completed high school	.41	.50	0	1
Age in 2008 in decades	3.78	1.40	1.5	7.6
Proportion of life in Alexandra	.66	.35	0	1
Socio-economic status	.59	.23	0	1
Employment: Unemployed	.30	.46	0	1
Employment: Not in the labour force	.20	.40	0	1
Housing: Shack	.25	.22	0	1
Housing: Hostel	.05	.44	0	1
Housing: Backyard room	.33	.47	0	1
Single	.32	.47	0	1
Zulu	.32	.47	0	1

Means and standard deviations are post-stratified.

Table A-3. Item Wording, Missing Values and Scale Construction

Participation in the 2008 attacks. For the next few questions, you will be filling in your own answers on a separate card to ensure your privacy. When I read out each question, you must mark your answers yourself on your card. For each question, you must make a cross. So if the 1st answer I read is the right one for you, then make a cross on ‘A’. If the 2nd answer is the right one for you, make a cross on ‘B’. If the 3rd answer is the right one, make a cross on ‘C’. Many people from Alex were involved in the attacks on foreigners in 2008. Can you tell me if any of the following things happened during those attacks? (I did this and feel that it was the right thing to do, I did this and regret it now = 1. I did not do this = 0). (1) Did you join in the toyi-toyi-ing and singing? (2) Did you threaten or intimidate anyone into joining in? (3) Did you steal any things from a foreigner’s shack? (4) Did you physically harm any foreigners? (5) Did you destroy anyone’s shack?

Missing values: 2 (0.4%)

Scale: Any affirmative response for (1) though (5) = 1; otherwise = 0.

Party support. Which of these political parties do you like the most? (DA, IFP, COPE, UDM, PAC, other = Opposition party. I don’t like any party, don’t know = None. ANC = ANC.)

Missing values: 43 (9.5%)

Socio-economic status. Do you or anyone else in your household own a ... (1) TV? (2) cell-phone? (3) fridge? (4) microwave? (5) computer? (6) car? (No = 1. Don’t know = missing. Yes = 0).

Missing values: 2–11 (0.4–2.4%)

Scale estimated using a 2-parameter logistic item response model. Scale reliability = .87

Employment status. Which one of these statements best describes your working life? (Working full-time, part-time, for myself = Working. Not working–housewife, student, retired, Unemployed–not looking for work = Not in the labor force. Unemployed–looking for work = Unemployed.)

Missing values: 0

Single. Are you married, in a relationship with someone, or single? (Married, in a relationship = 0. Single, divorced / widowed = 1).

Missing values: 5 (1.1%)

Meeting attendance. There are sometimes community meetings in Alex that are held by leaders such as the people from the community policing forum or the izinduna. So far this year, how many of these meetings in Alex have you attended? (0, don’t know = 0. 1 = 1. 2 or 3 = 2. 4 or more = 3).

Missing values: 1 (0.2%)

Education. What is the highest level of education you have personally achieved? (Primary school = 1. Some high school but no matric = 2. Finished matric, artisan’s certificate = 3. University degree, Teacher’s college diploma, Technikon diploma, Some other post-matric diploma = 4.)

Missing values: 1 (0.2%)

Proportion of life in Alexandra. (Number of years respondent has lived in Alexandra) / (Respondent’s age).

Missing values: 1 (0.2%)

3. Missing Data

Some respondents refused to provide answers for certain questions. Others responded with “don’t know.” Wherever possible, “don’t know” answers were recoded as some appropriate level on the response scale. Refusals were treated as missing values. Details on the extent of missingness by item and the treatment of “don’t know” responses are provided in Table A-3.

Missing values were handled using the method of multiple imputation. This procedure uses an algorithm to impute missing data based on the values of observed data (see Rubin 1987; King et al. 2001). A small number of complete datasets is created. Five is adequate for a sample of this size where most variables show less than 3% of their values missing (Rubin 1996). The algorithm is stochastic, such that each of the imputed datasets differs slightly from the others. Statistical analyses are then conducted on all data-sets before the results are combined. The additional uncertainty generated through the use of estimated data is reflected in the variance of the five imputed values for each missing datum, and ultimately in the variance of the parameter estimates in any model using multiply-imputed datasets.

In only one of the items used in this study were a significant proportion of the values missing: the question on party support. Given the history of partisan violence in Alexandra, it is not surprising that almost ten percent of respondents refused to offer an answer to this question. The issue is whether the missing values are missing at random (MAR) or not; in other words: if the missingness is uncorrelated with the values of the party variable itself. If so, multiple imputation provides an unbiased estimate of the true values. The MAR assumption would be violated in this case if supporters of certain parties were more likely to refuse to respond than others.

Further analysis of the pattern of missingness suggests that the MAR assumption is reasonable. A dummy variable for refusals for the party support question is correlated both with being Zulu (.24) and living in policing sector 2 (.30).¹³ Being Zulu in Alexandra is associated with being a supporter of a particular party, the IFP. This correlation suggests that missingness on this variable may be a function of being an IFP supporter. The geographic correlation suggests, on the other hand, that missingness is a function of where respondents live, not which parties they support. A logistic regression of the refusals dummy reveals that being Zulu is no longer a significant independent variable when living in Sector 2 is included in the equation. It is thus location, rather than ethnicity, which determines hesitance to answer the party support question. It thus seems likely that missing data on the party support question are MAR and thus amenable to multiple imputation.

¹³Sector 2 was the site of much of the partisan violence of the early 1990s.

4. Additional Results

Table A-4. Participation in 2008 by Riot Behaviour and Gender

Form of participation	Whole sample		Men		Women	
	<i>N</i>	% [95% C.I.]	<i>N</i>	% [95% C.I.]	<i>N</i>	% [95% C.I.]
Joined in the toyi-toying	40	7.8 [5.2 : 11.1]	28	11.2 [7.1 : 16.6]	12	4.1 [2.2 : 7.0]
Stole property	13	2.4 [1.0 : 5.0]	12	4.2 [1.5 : 8.9]	1	0.6 [0.0 : 3.4]
Used threats to coerce someone	11	1.5 [0.5 : 3.4]	11	3.0 [1.0 : 6.6]	0	0.0 —
Physical assault	17	3.3 [1.7 : 5.8]	12	4.3 [1.5 : 9.4]	5	2.2 [0.8 : 4.7]
Destroyed shacks	16	2.9 [1.2 : 5.9]	13	4.5 [1.6 : 9.7]	3	1.2 [0.1 : 4.8]
Any behaviour	44	8.5 [5.7 : 12.2]	29	11.6 [7.2 : 17.4]	15	5.3 [2.9 : 8.6]
Sample size	452		222		230	

N shows the number of respondents participating in each behavior and is unweighted. The percentages are post-stratified. Confidence intervals, in brackets, are design-based and utilize a beta binomial distribution.

Table A-5. Ordered Logit Regressions of Participation in 2008

	Whole sample		Men		Women	
Party support: none ^a	.70	(.50)	.80	(.58)	1.03	(.76)
Party support: opposition	1.38***	(.52)	2.08***	(.65)	1.15*	(.68)
Socio-economic status	-1.28	(1.10)	-2.27	(1.53)	-.64	(1.41)
Employment: unemployed ^b	.27	(.27)	-.17	(.54)	1.15*	(.62)
Employment: not in the labour force	1.36	(.98)	2.68**	(1.25)	-.10	(1.45)
Age (in decades)	3.09***	(1.17)	6.95***	(1.96)	-.70	(1.62)
Age (in decades) squared	-.42***	(.16)	-.90***	(.27)	.03	(.22)
Proportion of life spent in Alexandra	1.12	(.77)	1.59	(1.11)	.40	(1.29)
Housing: hostel ^c	1.08	(.81)	.32	(1.00)	2.63**	(1.22)
Housing: shack	.74	(.53)	.71	(.44)	.91	(1.02)
Housing: backyard room	.88	(.63)	1.31**	(.64)	-.37	(1.14)
Single	.01	(.31)	.45	(.37)	-.93	(.74)
Male	.75*	(.40)				
Meeting attendance	.32*	(.17)	.35	(.24)	.58**	(.27)
Education: primary school ^d	.51	(1.06)	.89	(1.36)	1.21	(1.91)
Education: some high school	.34	(.75)	.72	(1.08)	-.22	(1.55)
Education: completed high school	.98	(.64)	1.60**	(.81)	-.23	(1.62)
Language: Zulu	.02	(.40)	.27	(.55)	-.50	(.50)
<i>N</i>	452		222		230	

*** $p < .01$, ** $p < .05$, * $p < .10$. Outcome variable is coded “2” if respondent assaulted, intimidated, or destroyed property, “1” for other forms of participation, “0” otherwise. ^aOmitted category is ANC support. ^bOmitted category is employed. ^cOmitted category is formal house or flat. ^dOmitted category is higher education. Estimates are post-stratified. Design-based standard errors in parentheses.

Table A-6. Logit Regressions of Participation in 2008 With Missing Values Dropped

	Whole sample		Men		Women	
Party support: none ^a	.65	(.49)	.66	(.57)	1.00	(.69)
Party support: opposition	1.30***	(.49)	1.84***	(.57)	1.23*	(.67)
Socio-economic status	-1.27	(1.25)	-2.18	(1.73)	-.60	(1.46)
Employment: unemployed ^b	.27	(.28)	-.13	(.56)	1.21**	(.60)
Employment: not in the labour force	1.33	(1.06)	2.61*	(1.46)	-.08	(1.48)
Age (in decades)	2.80**	(1.21)	6.38***	(1.98)	-1.08	(1.44)
Age (in decades) squared	-.38**	(.16)	-.83***	(.28)	.08	(.19)
Proportion of life spent in Alexandra	1.19	(.79)	1.66	(1.12)	.43	(1.37)
Housing: hostel ^c	1.14	(.79)	.34	(1.04)	2.89**	(1.23)
Housing: shack	.81	(.54)	.78*	(.46)	1.06	(1.05)
Housing: backyard room	.90	(.63)	1.26**	(.63)	-.21	(1.18)
Single	.03	(.33)	.53	(.37)	-.97	(.80)
Male	.74*	(.39)				
Meeting attendance	.31*	(.17)	.34	(.24)	.56**	(.25)
Education: primary school ^d	.55	(1.03)	1.03	(1.26)	1.18	(1.92)
Education: some high school	.40	(.72)	.88	(.99)	-.24	(1.53)
Education: completed high school	1.07*	(.62)	1.74**	(.77)	-.15	(1.56)
Language: Zulu	-.02	(.39)	.22	(.51)	-.48	(.50)
Intercept	-9.91***	(2.20)	-17.15***	(3.29)	-0.90	(1.69)
<i>N</i>	450		220		230	

*** $p < .01$, ** $p < .05$, * $p < .10$. The two observations with missing values for the participation items are dropped. ^aOmitted category is ANC support. ^bOmitted category is employed. ^cOmitted category is formal house or flat. ^dOmitted category is higher education. Estimates are post-stratified. Design-based standard errors in parentheses.

References

- Afrobarometer Network. 2007. "Afrobarometer Round 4 Survey Manual." http://www.afrobarometer.org/files/documents/survey_manuals/survey_manual_r4_english.pdf:. Afrobarometer Network. January 25, 2013.
- Alexandra Renewal Project. ND. "Alex Benchmark Survey 2005." http://www.alexandra.co.za/downloads/oppres_benchmarksurvey_pres_0508.pdf: Alexandra Renewal Project. August 27, 2010.
- Bonner, Philip, and Noor Nieftagodien. 2008. *Alexandra: A History*. Johannesburg: Wits University Press.
- Dixon, John, and Clyde Tucker. 2010. "Survey Nonresponse." In *Handbook of Survey Research*, ed. Peter V. Marden, and James D. Wright. 2nd ed. London: Emerald pp. 593–630.
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1): 49–70.
- Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*. 16th ed. SIL International.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Rubin, Donald B. 1996. "Multiple Imputation After 18+ Years." *Journal of the American Statistical Association* 91(434): 473–89.
- Weisberg, Herbert F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago, IL: University of Chicago Press.
- Werner, Oswald, and Donald T. Campbell. 1970. "Translating, Working Through Interpreters, and the Problem of Decentering." In *A Handbook of Cultural Anthropology*, ed. Raoul Naroll, and Ronald Cohen. New York: American Museum of Natural History.