

Supplementary Material for:
Do Gender Quotas Really Reduce Bias? Evidence from a Policy
Experiment in Southern Africa

1 Explicit Bias: Regression Tables

DV: Political Bias	Model 1.1	Model 1.2	Model 1.3	Model 1.4
	Total	Total	Women	Men
(Intercept)	2.938 (0.058)	3.054 (0.048)	3.174 (0.068)	2.931 (0.064)
Quota	0.047 (0.108)	-0.086 (0.101)	-0.164 (0.160)	-0.007 (0.130)
Woman	0.227 (0.074)			
Quota x Woman	-0.068 (0.163)			
Under 25		-0.014 (0.105)	-0.043 (0.157)	0.026 (0.144)
Quota x Under 25		0.357 (0.180)	0.507 (0.264)	0.189 (0.280)
DV: Traditional Bias	Model 2.1	Model 2.2	Model 2.3	Model 2.4
	Total	Total	Women	Men
(Intercept)	2.386 (0.069)	2.527 (0.052)	2.662 (0.072)	2.390 (0.074)
Quota	0.043 (0.140)	0.115 (0.114)	0.085 (0.153)	0.143 (0.155)
Woman	0.337 (0.087)			
Quota x Woman	-0.090 (0.170)			
Under 25		0.112 (0.116)	0.267 (0.154)	-0.014 (0.164)
Quota x Under 25		-0.415 (0.221)	-0.514 (0.318)	-0.352 (0.291)
DV: Education Bias	Model 1	Model 2	Model 3	Model 4
	Total	Total	Women	Men
(Intercept)	3.432 (0.047)	3.507 (0.038)	3.573 (0.051)	3.440 (0.055)
Quota	0.052 (0.083)	0.006 (0.063)	0.027 (0.084)	-0.017 (0.103)
Woman	0.129 (0.061)			
Quota x Womoan	0.026 (0.108)			
Under 25		-0.044 (0.076)	-0.049 (0.116)	-0.032 (0.108)
Quota x Under 25		0.220 (0.102)	0.186 (0.137)	0.247 (0.179)

Table 1: Model-based ATE and CATE estimates. To correct for multiple sub-group tests, coefficients with $p < 0.0125$ are marked in **bold**. Robust standard errors clustered at the ED level. Num obs = 996, Num EDs = 135.

2 IAT Research Design

2.1 Field Sites

In February - March of 2014, I conducted 101 IATs in 92 field sites in Lesotho. I selected these sites from the complete list of electoral divisions in eight out of Lesotho's ten main administrative districts through R's random number generator. Because quota randomization was assigned at the level of the electoral division, I gained the most statistical power by maximizing the number of field sites I visited rather than the number of respondents per field site. Increasing the number of subjects per cluster has little effect on the treatment effect's standard error, as it does little to reduce the variance of the cluster-level outcomes (Gerber & Green, 2012: 83). I sampled in previously reserved and unreserved EDs evenly to have a similar number of observations in treatment and control conditions.

I conducted these tests with assistance from the Lesotho branch of a regional gender advocacy group, Gender Links, which receives funding from the Southern African Development Community (SADC). My two local research assistants from this organization, Alice Ranthimo and Thabiso Andrew Leche, provided invaluable translation assistance, in-depth local knowledge, and logistical support as I conducted these tests. The research assistants would explain the nature of the study to each respondent and ask whether they would like to participate in exchange for a small payment (20 South African Rand, equivalent to 2 USD). Participants took a basic gender-taste Brief Implicit Association Test (BIAT) as a warm up task to boost familiarity with the test's instructions prior to taking the main gender-occupation IAT. The main tests typically took between five to ten minutes to complete. To further ensure that the respondent was taking the test correctly, the test stopped automatically if the respondent averaged over six seconds in response time per prompt. If the respondent pressed the wrong key, he or she was

instructed to press the correct key; however, the test also stopped automatically if the respondent answered incorrectly to over 50 percent of associational tasks in any given block.

I exclude the two far eastern districts of Mokhotlong and Qacha's Nek from our sampling, which were difficult to access from the capital, Maseru. Given Lesotho's small size, I was able to conduct field tests in 80 percent of the country's geographic districts - but it should be noted that Mokhotlong and Qacha's Nek are particularly remote and mountainous areas. Whereas the data is a random sample across the whole country, the results from the implicit bias tests include a sample in which the lowland areas are somewhat overrepresented as compared to a whole county sample. However, given that the quota was randomly assigned within each of the ten large administrative districts, excluding these two districts does not threaten the internal validity of the research design or introduce bias when estimating the quota's treatment effect.

2.2 Statistical Power

To calculate the statistical power of the results, I simulate the minimum detectable coefficient for the conditional average treatment effects presented in Table 4. Table A1 displays there results. Despite the relatively small sample size, minimal variance in the outcome variable (the IAT's associated D score) lends credence to my ability to detect true average treatment effects.

	Total	Split Sample	Split Sample
	Respondents	Female	Male
Quota	0.101	0.095	0.101
I(Quota * Und. 25)	0.172	0.163	0.281
I(Quota * Female)	0.133	-	-

Table A1: Simulated power calculations for IAT regression results. Minimum detectable effect size reported.

2.3 Further description of IAT Test

I created the IATs through the program Inquisit by Millisecond Software. More information about the software can be found at: <http://www.millisecond.com/>. The IATs were adapted from those used by Beaman *et al.* (2009).

This description is adapted from Project Implicit at implicit.harvard.edu.

The IAT consists of five parts or blocks. The respondent starts with her fingers on the “e” and “i” keys of the laptop, which she must use to indicate whether the auditory or pictorial prompt is on the left (“e”) or right (“i”) side of the screen.

In the first block, the respondent sorts words relating to the concepts (e.g., leadership, domestic life) into categories. If the category “Leader” was on the left, when the respondent hears an authority prompt of a leadership word, she should press the “e” key on the left side of the screen.

In the second part of the IAT, the respondent sorts male and female Sesotho names with an image of a male or female Mosotho. So if the women’s image was on the right,

and the respondent heard a woman's name, she would press the "i" key.

In the third part of the IAT the categories are combined and the respondent is asked to sort both occupations and names to the appropriate column. So if the image categories on the left hand side are "Leader"/"Woman" then the categories on the right hand side will be "Home"/"Man". It is important to note that the order in which the blocks are presented varies across participants, so some respondents will do the "Leader"/"Woman", "Home"/"Man" block first and other people will do the "Leader"/"Man", "Home"/"Woman" block first.

In the fourth part of the IAT the placement of the concepts switches. If the category "Leader" was previously on the left, now it would be on the right. Importantly, the number of trials in this part of the IAT is increased in order to minimize the effects of practice.

In the final part of the IAT the categories are combined in a way that is opposite what they were before. If the category on the left was previously "Leader"/"Woman", it would now be "Leader"/"Man".

The IAT score is based on how long it takes the respondent, on average, to sort the words in the third part of the IAT versus the fifth part of the IAT. The respondent has an implicit association of men with leadership if they are faster to categorize words when Men and Leader share a response key and Women and Home share a response key, relative to the reverse.

Below is the list of words that the respondent hears to associate it with its respective appropriate image on either the left or right hand side of the screen. In Sesotho, the words in these categories are on average approximately the same length and one group of words does not take longer to read out than the other. The eight words associated respectively with the leader and domestic prompts, as well as the two main images associated with the domestic and leadership categories, were designed to be gender neutral - so that the associations remained with the larger concept and not with a concept that

has a further gendered association. For instance, I used the word “home” in the domestic category because it is gender-neutral, but not the word “cooking,” because it is a domestic task that is undertaken primarily by women.

Leadership Words (translated from Sesotho)

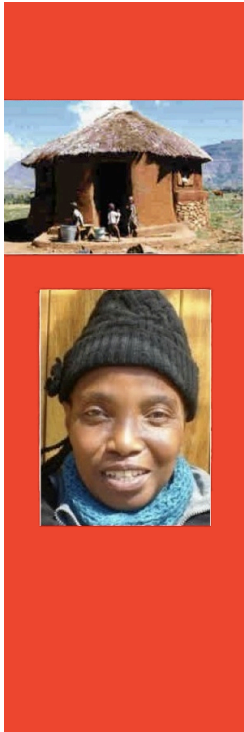
Councilor, public gathering (*pitso*), govern, chairperson, voting, leader, campaigning, director

Domestic Words (translated from Sesotho)

Home, parents, children, family, marriage, wedding, relatives, spouse

Gender-Occupation IAT screenshots

Stereotypical Block



“Lelapa” (“home” in Sesotho)
or
“Ntolo” (a female Sesotho name)

“Pitso” (“village meeting” in Sesotho)
or
“Tsepo” (a male Sesotho name)

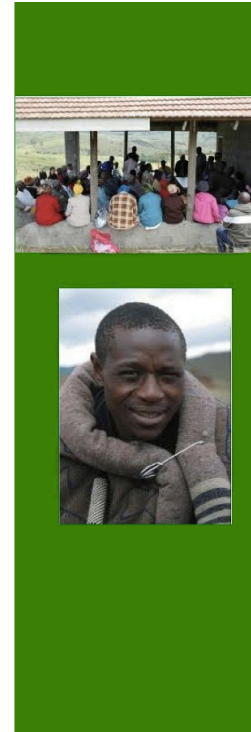
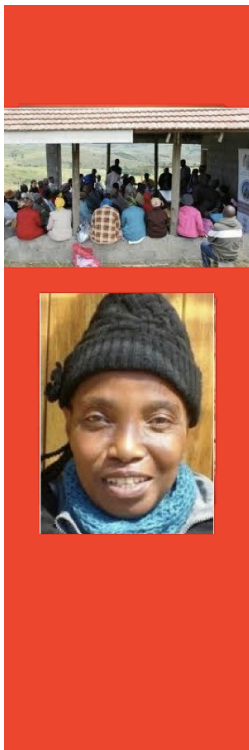


Figure A1: Sample screen shots from Gender-Occupation IAT (English translation): Stereotypical block (female/domestic and male/leader)

Counter- Stereotypical Block



“Pitso” (“village meeting” in Sesotho)
or
“Ntolo” (a female Sesotho name)

“Lelapa” (“home” in Sesotho)
or
“Tsepo” (a male Sesotho name)

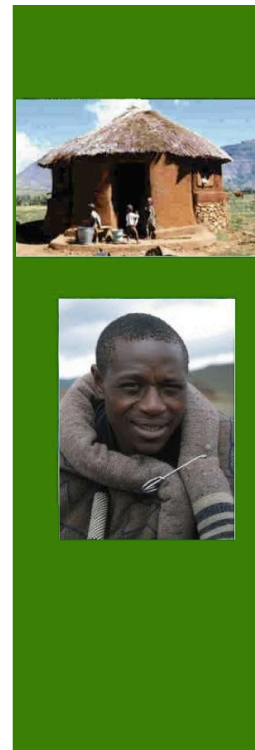


Figure A2: Sample screen shots from Gender-Occupation IAT (English translation): Counter-stereotypical block (female/leader and male/domestic)

IATs in the Field



Figure A3: Female respondent taking laptop-based IAT in the field.

	Model 1	Model 2	Model 3	Model 4
	Total	Total 2	Women 3	Men 4
(Intercept)	0.137	0.156	0.167	0.139
	(0.041)	(0.025)	(0.034)	(0.036)
Quota	0.010	0.021	0.045	0.002
	(0.049)	(0.032)	(0.042)	(0.046)
Woman	0.059			
	(0.052)			
Quota x Woman	-0.002			
	(0.063)			
Under 25		0.050	0.123	-0.006
		(0.069)	(0.060)	(0.103)
Quota x Under 25		-0.060	-0.160	0.026
		(0.080)	(0.072)	(0.118)
Num. obs.	101	101	51	50
Num. groups: ED	92	92	48	45

To correct for multiple sub-group tests, coefficients with $p < 0.0125$ are marked in **bold**

Table 2: D-score responses to IATS: Model-based estimates of age and gender CATEs. Robust standard errors clustered by ED.

3 Implicit Bias Tables

4 Balance Diagnostics

See Appendix Table A2 on next page

The only variable that achieves statistical significance is the average age of respondents: respondents from reserved EDs are on average 2.9 years older than in non-reserved EDs. As this is the only variable of 26 balance characteristics to achieve significance at the five-percent level, I attribute this difference to statistical chance. Further, a two-sample Kolmogorov-Smirnov test for equality of distributions is unable to reject that the two samples are drawn from the same population age distribution. As all other variables do not show significant differences between reserved and unreserved EDs, I conclude that the quota was indeed successfully randomized.

5 2012 Afro-Barometer Explicit Bias Indicators

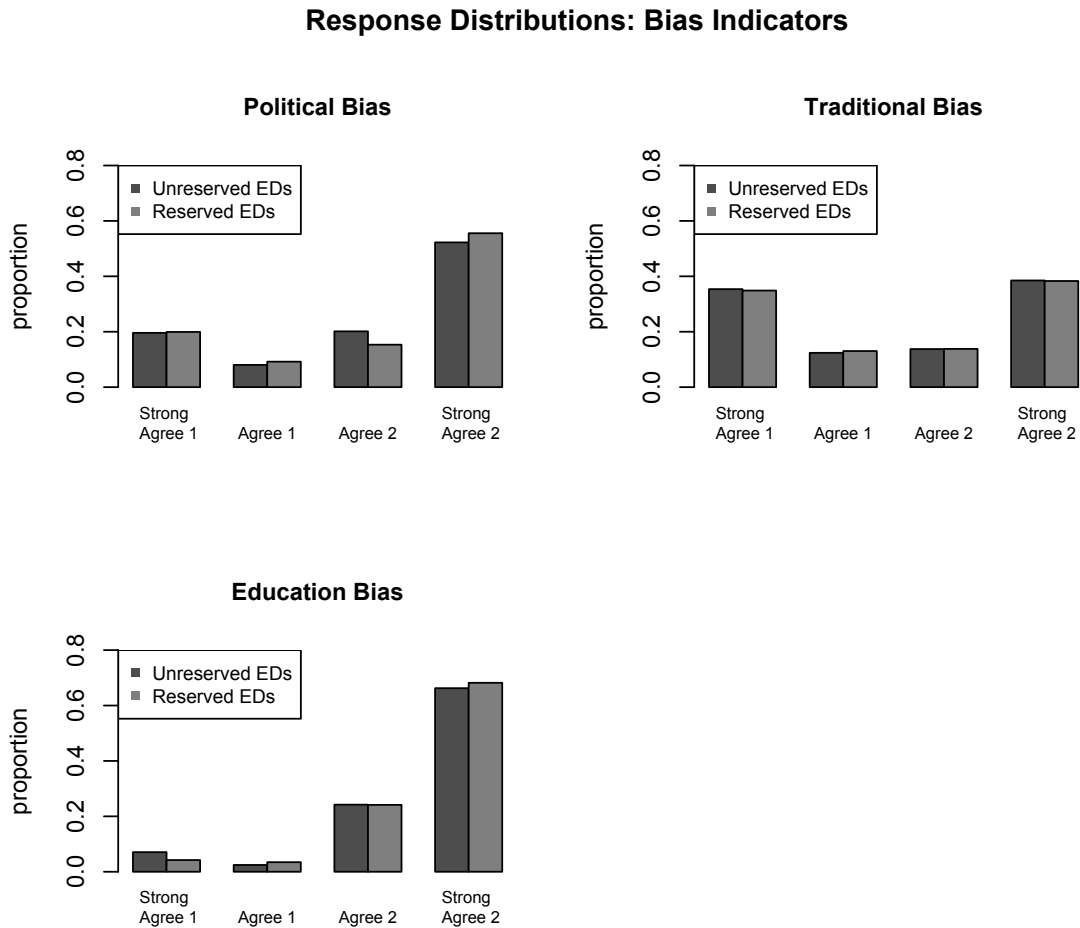


Figure A4: Three explicit bias indicators from 2012 Afro-barometer data: distribution of response categories in previously reserved versus unreserved electoral divisions.

References

- Beaman, Lori, Chattopadhyay, Raghavendra, Duflo, Esther, Pande, Rohini, & Topalova, Petia. 2009. Powerful Women: Does Exposure Reduce Bias? *Quarterly Journal of Economics*, **124**(4).
- Gerber, Alan S., & Green, Donald P. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York & London: W. W. Norton & Company, Inc.

	Reserved EDs		Unreserved EDs		Diff (Reserved- unreserved)	p-value	N
	Mean	SD	Mean	SD			
<i>Individual Characteristics</i>							
Age	45.395	(1.346)	42.49	(0.824)	2.905	(0.0331)	855
Children's age	5.049	(0.140)	2.144	(0.0829)	2.905	(0.443)	863
Gender (=1 if male)	0.4823	(0.0165)	0.508	(0.0107)	-0.0257	(0.123)	863
Self-reported poverty scale (1-10max)	2.715	(0.194)	2.506	(0.119)	0.209	(0.283)	853
Education level	2.77344	(0.191)	2.771	(0.0990)	0.00244	(0.990)	863
Religion	2.9464	(0.110)	2.932	(0.0609)	0.0144	(0.895)	863
Main Occupation	12.742	(2.247)	14.29	(1.499)	-1.548	(0.492)	862
Presence of cash income	0.2737	(0.0658)	0.320	(0.0403)	-0.0463	(0.484)	863
Agreement w/ government policies (1-5 (least agreement))	2.7326	(0.0938)	2.640	(0.0523)	0.0926	(0.326)	794
Member of a community self-help group (=1 if member)	0.2914	(0.0290)	0.265	(0.0161)	0.0264	(0.364)	863
Family working in South Africa (=1 if present)	0.3752	(0.0353)	0.352	(0.0203)	0.0232	(0.512)	859
Closeness to the ruling party	0.5148	(0.0445)	0.544	(0.0272)	-0.0292	(0.513)	863
<i>Village characteristics</i>							
Access to electricity grid (=1 if access)	0.2406	(0.0993)	0.192	(0.0475)	-0.0486	(0.625)	855
Access to piped water	0.2199	(0.103)	0.307	(0.0524)	0.0871	(0.398)	862
Access to sewage connection	0.1445	(0.0897)	0.0906	(0.0305)	-0.0539	(0.549)	862
Presence of health clinic in village	0.31	(0.0811)	0.163	(0.0378)	-0.147	(0.0728)	852
Presence of community meeting building in village	0.1549	(0.0868)	0.0986	(0.0303)	-0.0563	(0.518)	788
Presence of petrol station village	0.0649	(0.0368)	0.0325	(0.0160)	-0.0324	(0.380)	861
Presence of supermarket in village	0.1911	(0.0890)	0.221	(0.0413)	0.0299	(0.738)	863
Presence of school in village	0.60085	(0.0946)	0.595	(0.0555)	-0.00585	(0.951)	863
Presence of police station in village	0.0681	(0.0647)	0.128	(0.0371)	0.0599	(0.357)	863
Presence of bus/taxi stop in village	0.4354	(0.106)	0.397	(0.0554)	-0.0384	(0.719)	855
Presence of tarred road w/in 10km	0.5507	(0.106)	0.508	(0.0586)	-0.0427	(0.687)	863
Presence of graded/dirt/sand roads w/in 10km	0.62253	(0.0966)	0.631	(0.0542)	0.00847	(0.930)	863

Notes: 1.Data from the 2004 Afrobarometer survey.