Technical appendix

Strengthening accountability through media in Kenya: final evaluation

April 2017
Research and Learning

Contents

Introduction

This technical appendix is intended to be read alongside the associated BBC Media Action report, *Strengthening accountability through media in Kenya: final evaluation*. This is available from: http://dataportal.bbcmediaaction.org/site/assets/uploads/2017/04/Kenya-Country-Report-2017.pdf

1. Survey sampling methodology

As outlined in the table below, BBC Media Action carried out four national surveys.

Table 1: Quantitative surveys conducted (2012-2015)

Survey	Data collection	Sample size	Criteria
Baseline	August 2012	n=3,000	Nationally representative, adults 15+
Midline	June–July 2013	n=3,006	Nationally representative, adults 15+
Tracker (omnibus)	February 2014	n=2,031	Nationally representative, adults 18+
Endline	January–February 2015	n=3,003	Nationally representative, adults 15+

For each of these surveys, the sample was stratified across the major geographical divisions of the country (by province/region/state). Within these geographical divisions, a probability proportional to size multistage cluster sample was employed. At all stages, the selection of clusters was random and self-weighting. The sampling frame was constructed using the 2009 Kenya Population and Housing Census. Within enumeration areas, predefined random starting points were used to begin household selection. Random walk was applied with a fixed household interval. Within households, a Kish grid was used to select respondents.

Data collection was carried out using face-to-face interviews and recorded using either paper and pen, or computer-assisted personal interviewing (CAPI).

After data collection, the samples were compared to the latest census data and, where necessary to correct for any imbalances in region/state, gender, age and location (urban versus rural), nested weights applied.

2. Regression analysis

This technical appendix summarises the results of analysis BBC Media Action carried out on the Kenya midline (2013) and endline (2015) datasets. In both cases the samples were representative of Kenya's national adult population (aged 15+).

BBC Media Action carried out regression analysis to test the association between exposure to *Sema Kenya* (Kenya Speaks) and several governance outcomes (political knowledge, discussion, internal efficacy, external efficacy, participation and inclusive attitudes), while controlling for some of the other key factors which may influence these outcomes.

Variables

The independent variable for regression analysis was exposure to *Sema Kenya,* with two categories: regularly exposed to *Sema Kenya* (exposed to at least every other episode), and never exposed to *Sema Kenya*. Those who had been exposed, but not regularly, and those without access to media were set as missing.

The dependent variables were constructed as either categorical or continuous variables, dependent on the distribution of the outcome variables. Logistic regression was carried out for categorical dependent variables, and linear (OLS) regression was conducted for continuous dependent variables.

In addition to being based on past research and the specific country context, the confounders used in the analysis were chosen because they were hypothesised to be key factors in influencing the outcome variable. They therefore varied slightly across models.

Significance testing

Before carrying out regression analysis, BBC Media Action conducted statistical tests in order to measure the strength and the direction of bivariate relationships, as well as to test their significance. More precisely, BBC Media Action analysed:
- The relationship between the main independent variable (exposure) and the construct variables defined as outcomes (political knowledge, discussion, internal efficacy, external efficacy, participation and inclusive attitudes)
- The relationships among outcome variables
- The relationship between exposure and all the socio-demographic variables potentially associated with it (referred to as "confounders")
- The relationship between the outcome variables and confounders

BBC Media Action conducted different types of significance tests according to the nature of the variables considered. T-tests and Mann-Whitney U-tests were used to compare the differences between means, Pearson's R and Spearman's Rho tests were used to ascertain correlation, and Chi-squared tests were conducted to measure associations. All tests were conducted with significance at the $p = 0.05$ level.

Analysis

As mentioned above, BBC Media Action carried out different types of regression analysis based on the dependent variable.

> **Logistic regression model:** this allows researchers to work with categorical variables such as the binary variables where the distribution of the variables does not follow a normal and linear distribution that could have fitted better in another statistical model such as a linear regression. The logistic regression produces a probability value or odds ratio (OR)

that indicates how much more likely it is that cases with specific attributes will fit into a model that explains the presence of certain outcomes. The regressions are calculated with a certain degree of confidence specified by the model. This confidence interval is used to understand if the changes in one variable are associated with changes in the other as a result of a statistical relationship that can be explained by the model. Here, any value above 95% is considered as statistically significant.

The ordinary least squares (OLS) model: this allows researchers to work with a continuous dependent variable, derived through confirmatory factor analysis, and independent variables that have either continuous or categorical values. The regression coefficient for the independent variable provides key information indicating the estimated change in the dependent variable associated with a one unit increase in the independent variable. The model seeks to summarise this association by fitting a straight line to predict the value of the dependent variable based on the observed values of the independent variables.

BBC Media Action's data satisfied the principle assumptions required for justifying the use of OLS: the relationships between the dependent and independent variables were linear and additive, and the error terms were normally distributed, constant, and were not correlated. With these assumptions met, a confidence interval for the regression line was calculated for each estimate and BBC Media Action was able to test whether the hypothesis of a zero slope – that is of no relationship between the two key variables of interest – existed in the true population.

Prior to analysis, BBC Media Action adopted the conventional standard of rejecting the null hypothesis at the 0.05 level. Given this, BBC Media Action expects that any estimated effects that are significantly associated with exposure to the programme of interest fall within the range reported in the confidence intervals 95% of the time.

Table 2: Overview of regression models

Model	Dataset	Model performance			Association with exposure		
		Sample size	R square	Significance (OLS only)	Association (OR/ coefficient)	95% confidence interval/ standard error (SE)	Significance
Regression 1.1 Substantial knowledge versus low knowledge logistic model	Midline	999	0.537	-	4.532 (OR)	SE 0.346	0.001
Regression 1.2 Moderate versus low knowledge logistic model	Midline	1,491	0.196	-	2.009 (OR)	SE 0.294	0.030
Regression 1.3	Midline	1,748	0.161	-	1.529 (OR)	SE 0.158	0.009

Substantial versus moderate knowledge logistic model							
Regression 1.4 Knowledge multinomial model	Midline	2,245	0.279	<0.001	Moderate 2.101 (OR), Substantial 3.295 (OR)	Moderate SE 0.293, Substantial SE 0.301	Moderate 0.017, Substantial 0.001
Regression 1.5 Knowledge OLS model	Endline	2,223	0.186	<0.001	0.696 (B)	3.079-3.964	<0.001
Regression 2.1 Frequent discussion versus occasional discussion logistic model	Endline	1,146	0.178	-	2.018 (OR)	SE 0.212	0.004
Regression 2.2 Discussion linear model with interaction effects exposure x gender	Endline	2,233	0.165	<0.001	Exposure for women 0.749 (B)	0.034-1.464	0.040
Regression 3.1 High internal efficacy versus low internal efficacy logistic model	Endline	519	0.515	-	2.958	SE 1.219	0.026
Regression 4.1 High external efficacy versus low external efficacy logistic model	Endline	816	0.230	-	1.094 (OR)	SE 0.267	0.705
Regression 4.2 Medium external efficacy versus	Endline	1,017	0.105	-	1.317 (OR)	SE 0.261	0.443

low external efficacy logistic model							
Regression 4.3 High external efficacy versus medium external efficacy logistic model	Endline	2,134	0.108	-	0.867 (OR)	SE 0.205	0.607
Regression 5.1 Participated at least once versus never participated logistic model	Endline	2,233	0.180	<0.001	1.962 (OR)	0.939-4.100	0.018
Regression 5.2 Participation OLS model	Endline	2,233	0.284	<0.001	0.355 (B)	0.100-0.610	0.006
Regression 6.1 High versus moderate inclusiveness logistic model	Endline	2,423	0.078	-	2.611	SE 0.397	0.016

Full model results

Note: 'Ref' indicates the reference category of each variable

Knowledge

Table 3: Regression 1.1 – Substantial knowledge versus low knowledge logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	.346	.001	4.532
Male	Ref	-	-
Female	.197	.001	.366
Education - no schooling	Ref	-	-
Some primary	.485	.087	1.879
Completed primary	.437	.13	2.610
Completed secondary	.435	.004	3.309
College/university	.542	.001	7.103
Income – cannot afford food	Ref	-	-
Can afford food but purchasing clothes is a serious problem	.400	.004	2.360
Can afford food and clothes but purchasing of durables is difficult	.368	.001	3.334
Can afford main household appliances but purchasing a car is beyond our means	.415	.001	3.745
Earnings sufficient to buy anything except expensive purchases such as a house	.657	.010	3.967
Don't face financial problems	.6.665	.200	1.714
Internet – no access	Ref	-	-
Over a year ago	1.873	.250	2.463
Within the last year	1.067	.021	3.627
Within the last month	1.310	.004	6.421
Within the last week	.319	.932	1.034
Yesterday or today	.380	.001	3.463
Group membership - non-member	Ref	-	-
Inactive member	.353	.424	1.365
Active member	.281	.001	3.435

Official leader	.419	.001	7.868
Interest in politics - not at all interested	Ref	-	
Not very interested	.371	.002	3.572
Somewhat interested	.350	.001	13.044
Very interested	.350	.001	21.521
Constant	.636	.001	.009

The Nagelkerke R statistic for this model was .537. The Hosmer and Lemeshow statistic had a chi-square of 10.321 and a significance level of .243.

Table 4: Regression 1.2 – Moderate versus low knowledge logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	.294	.030	2.009
Male	Ref	-	-
Female	.141	.001	.620
Internet – no access	Ref	-	-
Over a year ago	1.925	.309	1.815
Within the last year	.460	.025	2.459
Within the last month	1.638	.006	5.409
Within the last week	.229	.558	1.186
Yesterday or today	.255	.001	2.528
Group membership - non-member	Ref	-	-
Inactive member	.222	.024	1.683
Active member	.177	.001	2.791
Official leader	.286	.001	4.308
Interest in politics - not at all interested	Ref	-	-
Not very interested	.179	.001	2.655
Somewhat interested	.195	.001	5.316
Very interested	.219	.001	4.750
Constant	.237	.002	.441

The Nagelkerke R statistic here was .196. The Hosmer and Lemeshow statistic had a chi-square of 8.685 and a significance level of .370.

Table 5: Regression 1.3 – Substantial versus moderate knowledge logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	.158	.009	1.529
Male	Ref	-	-
Female	.108	.001	.609
Internet – no access	Ref	-	-
Over a year ago	.344	.288	1.461
Within the last year	.258	.053	1.696
Within the last month	.229	.013	1.793
Within the last week	.177	.050	1.346
Yesterday or today	.146	.001	1.843
Group membership - non-member	Ref	-	-
Inactive member	.246	.272	1.364
Active member	.202	.027	1.575
Official leader	.240	.002	2.200
Interest in politics - not at all interested	Ref	-	-
Not very interested	.295	.857	1.108
Somewhat interested	.277	.007	2.076
Very interested	.280	.001	4.204
Constant	.330	.001	0.147

The Nagelkerke R statistic for this model was .161. The Hosmer and Lemeshow test had a chi-square statistic of 5.176 and a significance level of .739.

Table 6: Regression 1.4 – Knowledge multinomial model
In the multinomial models for knowledge, regular exposure to *Sema Kenya* was significantly associated with all levels of knowledge.

Moderate knowledge	Variable	Standard error	Significance level	Odds ratio
	Not exposed to *Sema Kenya*	Ref	-	-
	Regularly exposed	.294	.017	2.101

	to *Sema Kenya*			
	Female	Ref	-	-
	Male	.137	.001	1.582
	Education - college/university	Ref	-	-
	No schooling	.415	.001	.317
	Some primary	.350	.003	.340
	Completed primary	.307	.006	.459
	Completed secondary	.296	.043	.581
	Internet – over a year ago	Ref	-	-
	No access	1.258	.464	.613
	Yesterday or today	1.302	.782	1.116
	Within the last week	1.287	.355	.557
	Within the last month	2.229	.170	2.609
	Within the last year	1.571	.665	1.249
	Group membership – official leader	Ref	-	-
	Non-member	.294	.001	.234
	Inactive member	.299	.001	.368
	Active member	.260	.112	.613
	Interest in politics – very interested	Ref	-	-
	Not at all interested	.211	.001	.219
	Not very interested	.185	.001	.591
	Somewhat interested	.190	.401	1.139
Substantial knowledge				
	Not exposed to *Sema Kenya*	Ref	-	-
	Regularly exposed to *Sema Kenya*	.301	.001	3.295
	Female	Ref	-	-
	Male	.153	.001	2.583
	Education -	Ref	-	-

college/university			
No schooling	.511	.001	.112
Some primary	.373	.001	.219
Completed primary	.325	.001	.337
Completed secondary	.304	.004	.419
Internet – over a year ago	Ref	-	-
No access	1.284	.185	.421
Yesterday or today	1.316	.669	1.194
Within the last week	1.315	.272	.453
Within the last month	2.247	.126	3.049
Within the last year	1.595	.630	1.346
Group membership – official leader	Ref	-	-
Non-member	.331	.001	.115
Inactive member	.326	.001	.233
Active member	.282	.012	.454
Interest in politics – very interested	Ref	-	-
Not at all interested	.302	.001	.052
Not very interested	.214	.001	.156
Somewhat interested	.201	.015	.563

The Nagelkerke R statistic for this model was .279 and the Pearson chi-square had a value of 1602.884 and a significance level of .000.

Table 7: Regression 1.5 – Knowledge OLS model

Dependent variable: average self-reported knowledge (0 to 10)

Predictor	Unstandardized coefficients		Standardized coefficient	Significance	95.0% confidence interval for B	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sema Kenya*	Ref	-	-	-	-	-
Regularly exposed to	0.696	0.157	0.086	0.000	0.387	1.004

Sema Kenya						
Group membership – not an active member	Ref	-	-	-	-	-
Active member	0.538	0.123	0.086	0.000	0.297	0.778
Interest in politics - not at all interested	Ref	-	-	-	-	-
Not very interested	0.724	0.159	0.119	0.000	0.412	1.035
Somewhat interested	1.144	0.148	0.215	0.000	0.854	1.434
Very interested	1.809	0.15	0.35	0.000	1.516	2.102
Region – Rift Valley	Ref	-	-	-	-	-
Central	0.379	0.162	0.052	0.019	0.062	0.696
Coast	0.097	0.189	0.011	0.609	-0.273	0.467
Eastern	-0.217	0.153	-0.032	0.156	-0.517	0.083
Nairobi	-0.481	0.193	-0.059	0.013	-0.86	-0.102
North Eastern	1.463	0.242	0.13	0.000	0.988	1.938
Nyanza	-0.252	0.16	-0.035	0.116	-0.566	0.062
Western	-0.217	0.178	-0.027	0.224	-0.567	0.133
Male	Ref	-	-	-	-	-
Female	-0.422	0.102	-0.086	0.000	-0.621	-0.222
Rural	Ref	-	-	-	-	-
Urban	0.03	0.111	0.006	0.785	-0.188	0.248
Age 15-24	Ref					
Age 25-34	0.19	0.125	0.037	0.130	-0.056	0.436
Age 35-44	0.428	0.16	0.066	0.008	0.114	0.741
Age 45-54	0.513	0.202	0.058	0.011	0.118	0.909
Age 55-64	0.474	0.267	0.038	0.076	-0.05	0.999
Age 65+	0.611	0.32	0.04	0.056	-0.016	1.238
Education – illiterate	Ref	-	-	-	-	-
Literate	-0.688	0.163	-0.091	0.000	-1.008	-0.368
Completed primary	-0.105	0.12	-0.019	0.382	-0.342	0.131
Completed college/university	0.65	0.13	0.108	0.000	0.394	0.905
Income - medium	Ref	-	-	-	-	-
Low	0.105	0.118	0.019	0.375	-0.127	0.337
High	0.248	0.188	0.026	0.185	-0.119	0.616
Marital status – married, living with spouse	Ref	-	-	-	-	-
Single	0.075	0.125	0.015	0.549	-0.17	0.32
Married, not living with spouse	-0.174	0.16	-0.022	0.279	-0.488	0.141
Divorced/separated	-0.179	0.297	-0.012	0.547	-0.762	0.404
Widowed	-0.265	0.253	-0.022	0.294	-0.761	0.231
In a marriage where the husband has more than one wife	-0.406	0.713	-0.011	0.570	-1.805	0.993
Living with partner	-1.748	1.041	-0.032	0.093	-3.788	0.293
Constant	3.521	0.226		0.000	3.079	3.964

The model had an adjusted R square of 0.186. The Durbin-Watson value was 1.938. The F statistic was 18.252 (significance < 0.001).

Discussion

Note: there is no multinomial model for discussion, because only one model (frequent versus occasional) was tested. The 'never discussed' category held too few respondents to be able to run a meaningful regression.

Table 8: Regression 2.1 – Frequent discussion versus occasional discussion logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	0.212	0.004	2.018
Education – none or some primary	Ref	0.066	-
Completed primary	0.090	0.003	2.522
Completed secondary	0.447	0.070	1.566
Completed college/university	0.450	0.020	1.822
Region – Central	Ref	0.000	-
Coast	0.306	0.445	0.747
Eastern	0.241	0.023	1.574
Nairobi	0.242	0.051	0.635
North Eastern	6.685	0.017	5.581
Nyanza	0.243	0.407	0.747
Rift Valley	0.212	0.001	0.461
Western	0.317	0.949	1.080
TV – every week	Ref	0.054	-
At least twice a month	0.281	0.235	0.808
At least once a month	0.413	0.075	2.320
No access or less than once a month	0.427	0.314	0.614
Internet – every week	Ref	0.075	-
At least twice a month	0.227	0.071	0.650
At least once a month	0.243	0.387	0.778
No access or less than once a month	0.281	0.193	1.366
Newspaper – every week	Ref	0.001	-
At least twice a month	0.166	0.001	0.530
At least once a month	0.201	0.016	0.609
No access or less than once a month	0.271	0.245	0.791

Perceived freedom – none	Ref	0.103	-
Some	0.149	0.785	1.034
Substantial	0.166	0.119	1.394
Interest in politics - not at all interested	Ref	0.000	-
Not very interested	0.275	0.460	1.294
Somewhat interested	0.266	0.004	1.981
Very interested	0.264	0.001	3.091
Constant	0.510	0.088	0.560

The Nagelkerke R statistic for this model was 0.178. The Hosmer and Lemeshow statistic had a chi-square of 4.318 and a significance level of 0.827.

Table 9: Regression 2.2 – Discussion linear model with interaction effects exposure x gender

Dependent variable: average discussion frequency (0 to 10)

Predictor	**Unstandardized coefficients**		**Standardized coefficient**	**Significance**	**95.0% confidence interval for B**	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sema Kenya*	Ref	-	-	-	-	-
Regularly exposed to *Sema Kenya*	.614	.211	.072	.004	.202	1.027
Male	Ref	-	-	-	-	-
Female	-.567	.114	-.110	.000	-.791	-.343
Exposure for Women	.749	.365	.051	.040	.034	1.464
Group membership – not an active member	Ref	-	-	-	-	-
Active member	.298	.131	.045	.023	.041	.555
Interest in politics - not at all interested	Ref	-	-	-	-	-
Not very interested	.633	.170	.099	.000	.300	.966
Somewhat interested	1.202	.158	.214	.000	.892	1.513
Very interested	1.802	.160	.330	.000	1.489	2.116
Region – Rift Valley	Ref	-	-	-	-	-
Central	.499	.173	.065	.004	.160	.838
Coast	-.272	.202	-.030	.178	-.668	.124
Eastern	.976	.164	.137	.000	.655	1.296

Nairobi	.147	.207	.017	.479	-.259	.552
North Eastern	1.859	.259	.156	.000	1.350	2.367
Nyanza	-.005	.172	-.001	.976	-.341	.331
Western	.484	.191	.056	.011	.110	.858
Rural	Ref	-	-	-	-	-
Urban	.061	.119	.011	.609	-.172	.294
Age 15-24	Ref	-	-	-	-	-
Age 25-34	.182	.134	.033	.176	-.081	.445
Age 35-44	.333	.171	.049	.051	-.002	.669
Age 45-54	.608	.216	.065	.005	.185	1.032
Age 55-64	.583	.286	.044	.041	.023	1.144
Age 65+	.309	.342	.019	.367	-.362	.979
Education - Illiterate	Ref	-	-	-	-	-
Literate	-.044	.175	-.005	.803	-.386	.299
Completed primary	-.020	.129	-.003	.878	-.272	.233
Completed college/university	.080	.139	.013	.566	-.194	.353
Income - Medium	Ref	-	-	-	-	-
Low	-.403	.127	-.068	.001	-.651	-.155
High	.014	.201	.001	.943	-.379	.408
Marital status – married, living with spouse	Ref	-	-	-	-	-
Single	-.132	.134	-.025	.324	-.395	.131
Married, not living with spouse	-.333	.172	-.040	.052	-.669	.004
Divorced/separated	-.562	.318	-.035	.077	-1.186	.062
Widowed	-.515	.271	-.040	.057	-1.046	.016
In a marriage where the husband has more than one wife	.461	.764	.012	.546	-1.037	1.958
Living with partner	-.602	1.113	-.010	.589	-2.785	1.581
Constant	3.684	.242		.000	3.209	4.158

The model had an adjusted R square of 0.165. The Durbin-Watson value was 1.978. The F statistic was 15.360 (significance < 0.001).

Internal efficacy

Table 10: Regression 3.1 – High internal efficacy versus low internal efficacy logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	1.219	0.026	2.958
Male	Ref	-	-
Female	0.279	0.001	0.385
Region - Central	Ref	0.030	-
Coast	0.672	0.740	1.346
Eastern and North Eastern	0.492	0.053	2.786
Nairobi	0.551	0.278	0.521
Nyanza	0.508	0.602	0.770
Rift Valley	0.465	0.704	0.869
Western	0.635	0.583	0.821
Urban	Ref	-	-
Rural	0.332	0.024	0.535
Perceived identity - Kenyan only	Ref	0.039	-
More Kenyan than ethnic group	0.292	0.027	1.873
Neither Kenyan nor ethnic group	0.379	0.803	0.906
More ethnic group than Kenyan	1.704	0.032	4.060
Ethnic group only	4.889	0.360	1.965
Internet - no access	Ref	0.003	-
Every week	0.673	0.257	2.308
At least twice a month	0.756	0.617	0.855
At least once a month	0.784	0.327	0.563
Less than once a month	0.796	0.475	0.851
Perceived freedom - limited	Ref	0.016	-
Some	0.317	0.106	1.739
Substantial	0.382	0.016	2.517
Interest in politics - not at all interested	Ref	0.000	-
Not very interested	0.453	0.118	2.533
Somewhat interested	0.428	0.001	5.197

| Very interested | 0.464 | 0.001 | 32.700 |
| Constant | 0.905 | 0.067 | 0.145 |

The Nagelkerke R statistic for this model was 0.515. The Hosmer and Lemeshow statistic had a chi-square of 7.662 and a significance level of 0.467.

External efficacy

Table 11: Regression 4.1 – High external efficacy versus low external efficacy logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly reached by *Sema Kenya*	0.267	0.705	1.094
Education – no schooling	Ref	0.000	-
Completed primary	0.304	0.778	1.214
Completed secondary	0.302	0.003	0.623
Completed college/university	0.331	0.001	0.326
Region - Central	Ref	0.000	-
Coast	0.371	0.212	0.633
Eastern	0.301	0.001	0.364
Nairobi	0.336	0.002	0.323
North Eastern	3.565	0.451	2.356
Nyanza	0.339	0.001	0.192
Rift Valley	0.311	0.001	0.345
Western	0.381	0.001	0.258
TV – every week	Ref	0.001	-
Twice a month	0.333	0.001	2.179
Once a month	0.335	0.104	1.940
No access or less than once a month	0.328	0.004	2.399
Perceived freedom - none	Ref	0.000	-
Some	0.183	0.000	2.096
Substantial	0.209	0.000	2.907
Constant	0.393	0.018	1.823

The Nagelkerke R statistic for this model was 0.230. The Hosmer and Lemeshow statistic had a chi-square of 5.169 and a significance level of 0.739.

Table 12: Regression 4.2 – Medium external efficacy versus low external efficacy logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly reached by *Sema Kenya*	0.261	0.443	1.317
Education – none or some primary[1]	Ref	0.552	-
Completed primary	1.578	0.930	0.952
Completed secondary	1.572	0.764	0.738
Completed college/university	1.579	0.539	0.628
Income – cannot afford food	Ref	0.053	
Can afford food but purchasing clothes is a serious problem	1.423	0.632	0.862
Can afford food and clothes but purchasing of durables is difficult	1.394	0.944	0.647
Can afford main household appliances but purchasing a car is beyond our means	1.384	0.877	0.697
Earnings sufficient to buy anything except expensive purchases such as a house	1.399	0.698	0.524
Don't face financial problems	1.456	0.025	0.159
Region - Central[2]	Ref	0.140	-
Coast	0.451	0.181	0.664
Eastern	0.331	0.047	0.576
Nairobi	0.336	0.045	0.527
North Eastern	9.422	0.318	0.595
Nyanza	0.338	0.167	0.636
Rift Valley	0.302	0.527	0.892
Western	0.409	0.005	0.389
Perceived identity - Kenyan only	Ref	0.024	-
More Kenyan than ethnic group	0.203	0.345	1.147
Neither Kenyan nor ethnic group	0.251	0.939	1.014

---

[1] In this model, education was not significant at any level using the Wald statistic but was significant according to the Likelihood Ratio. As education was considered an important confounder, BBC Media Action kept it in the model.

[2] In this model, region was not significant at any level using the Wald statistic but was significant according to the Likelihood Ratio. As region was considered an important confounder, BBC Media Action kept it in the model.

More ethnic group than Kenyan	0.366	0.012	0.455
Ethnic group only	0.587	0.011	0.323
Watch TV – no access or less than once a month[3]	Ref	0.289	-
Every week	6.498	0.055	0.375
At least twice a month	6.539	0.394	0.655
At least once a month	6.531	0.038	0.318
Internet – no access	Ref	0.779	-
Every week	3.951	0.567	0.743
At least twice a month	3.937	0.679	0.821
At least once a month	3.961	0.846	1.061
Less than once a month	3.987	0.852	1.122
Newspaper – no access	Ref	0.009	-
Every week	6.565	0.663	0.643
At least twice a month	6.559	0.690	1.036
At least once a month	6.577	0.540	1.202
Less than once a month	6.558	0.151	0.346
Constant	10.588	0.003	49.756

The Nagelkerke R statistic for this model was 0.105. The Hosmer and Lemeshow statistic had a chi-square of 1.907 and a significance level of 0.984.

Table 13: Regression 4.3 – High external efficacy versus medium external efficacy logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	0.205	0.607	0.867
Education – no schooling	Ref	0.000	-
Some primary	0.273	0.955	1.056
Completed primary	0.253	0.599	1.031
Completed secondary	0.244	0.018	0.731
Completed college/university	0.279	0.002	0.376
Region - Central	Ref	0.000	-

---

[3] Higher standard error for TV, internet and newspapers might be due to colinearity, though original testing did not show any. Pearson's R ranged from 0.156 to 0.284 and was significant at p=0.000.

Coast	0.229	0.030	1.692
Eastern	0.191	0.894	0.979
Nairobi	0.254	0.910	1.138
North Eastern	0.260	0.049	1.770
Nyanza	0.247	0.005	0.493
Rift Valley	0.189	0.004	0.562
Western	0.230	0.789	0.911
Perceived freedom - none	Ref	0.002	-
Some	0.137	0.015	1.333
Substantial	0.145	0.002	1.684
Group membership –non-member		0.048	Ref
Inactive member	0.266	0.165	0.695
Active member	0.186	0.408	0.91
Official leader	0.218	0.205	1.267
Interest in politics – not at all interested	Ref	0.003	-
Not very interested	0.203	0.949	1.066
Somewhat interested	0.184	0.242	0.763
Very interested	0.177	0.059	1.287
Constant	10.588	0.003	49.756

The Nagelkerke R statistic for this model was 0.108. The Hosmer and Lemeshow statistic had a chi-square of 8.431 and a significance level of 0.393.

The relationship between external efficacy and regularly watching *Sema Kenya* was insignificant at all levels, indicating that exposure to the programme did not appear to lead to an increased belief in the capabilities of the government to serve the Kenyan people. Associations found between exposure and external efficacy may have happened by chance.

Participation

Table 14: Regression 5.1 – Participated at least once versus never participated logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	.286	.018	1.962
Group membership – not an active member	Ref	-	-
Active member	.145	.000	2.096
Interest in politics – not at all interested	Ref	.000	-
Not very interested	.191	.739	1.066
Somewhat interested	.187	.007	1.655
Very interested	.198	.001	1.949
Male	Ref	-	-
Female	.136	.000	.470
Rural	Ref	-	-
Urban	.131	.962	1.006
Age 15-24	Ref	.000	-
Age 25-34	.155	.000	1.835
Age 35-44	.227	.000	2.695
Age 45-54	.393	.000	5.033
Age 55-64	.743	.008	7.186
Age 65+	1.031	.042	8.116
Education – illiterate	Ref	.005	-
Literate, but no schooling	.782	.582	.650
Completed primary	.767	.540	.625
Completed secondary	.768	.771	.800
Completed college/university	.784	.676	1.388
Income - medium	Ref	.132	-
Low	.175	.104	1.330
High	.281	.186	1.450
Marital status – married, living with spouse	Ref	.156	-
Single	.161	.027	.700

	.220	.178	.744
Married, not living with spouse	.220	.178	.744
Divorced/separated	.552	.343	1.688
Widowed	.624	.206	2.202
In a marriage where the husband has more than one wife	1.135	.960	.944
Living with partner	.978	.475	.497
Constant	.798	.217	2.679

The Nagelkerke R statistic for this model was 0.180. The Hosmer and Lemeshow statistic had a chi-square of 6.169 and a significance level of 0.628.

Table 15: Regression 5.2 – Participation OLS model

Dependent variable: average political participation, including voting (0 to 10)

Predictor	Unstandardized coefficients		Standardized coefficient	Significance	95.0% confidence interval for B	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sema Kenya*	Ref	-	-	-	-	-
Regularly exposed to *Sema Kenya*	0.355	0.13	0.05	0.006	0.1	0.61
Group membership – not an active member	Ref	-	-	-	-	-
Active member	0.706	0.101	0.127	0.000	0.507	0.904
Interest in politics - not at all interested	Ref	-	-	-	-	-
Not very interested	0.169	0.131	0.032	0.198	-0.088	0.427
Somewhat interested	0.519	0.122	0.111	0.000	0.279	0.759
Very interested	0.947	0.124	0.208	0.000	0.705	1.19
Region – Rift Valley	Ref	-	-	-	-	-
Central	-0.115	0.134	-0.018	0.388	-0.378	0.147
Coast	-0.593	0.156	-0.079	0.000	-0.899	-0.287
Eastern	0.11	0.126	0.018	0.385	-0.138	0.358
Nairobi	-0.451	0.16	-0.063	0.005	-0.765	-0.138
North Eastern	1.507	0.2	0.151	0.000	1.114	1.9
Nyanza	-0.239	0.133	-0.037	0.072	-0.499	0.021
Western	-0.115	0.147	-0.016	0.435	-0.404	0.174
Male	Ref	-	-	-	-	-
Female	-0.837	0.084	-0.194	0.000	-1.001	-0.672
Rural	Ref	-	-	-	-	-
Urban	-0.075	0.092	-0.017	0.412	-0.255	0.105
Age 15-24	Ref	-	-	-	-	-
Age 25-34	0.608	0.104	0.133	0.000	0.404	0.811
Age 35-44	1.083	0.132	0.189	0.000	0.824	1.343
Age 45-54	1.435	0.167	0.184	0.000	1.107	1.762

Age 55-64	1.685	0.221	0.154	0.000	1.252	2.118
Age 65+	1.784	0.264	0.134	0.000	1.265	2.302
Education - illiterate	Ref	-	-	-	-	-
Literate	0.018	0.135	0.003	0.897	-0.247	0.282
Completed primary	-0.01	0.1	-0.002	0.923	-0.205	0.186
Completed college/university	0.359	0.108	0.068	0.001	0.148	0.57
Income – medium	Ref	-	-	-	-	-
Low	0.075	0.098	0.015	0.445	-0.117	0.266
High	0.168	0.155	0.02	0.280	-0.137	0.472
Marital status – married, living with spouse	Ref	-	-	-	-	-
Single	-0.494	0.103	-0.111	0.000	-0.697	-0.291
Married, not living with spouse	-0.26	0.133	-0.037	0.050	-0.52	0
Divorced/separated	-0.096	0.246	-0.007	0.697	-0.578	0.386
Widowed	-0.172	0.209	-0.016	0.411	-0.582	0.238
In a marriage where the husband has more than one wife	-0.701	0.59	-0.021	0.235	-1.857	0.456
Living with partner	0.022	0.86	0	0.980	-1.665	1.709
Constant	3.473	0.187		0.000	3.107	3.838

The model had an adjusted R square of 0.284. The Durbin-Watson value was 1.931. The F statistic was 30.808 (significance < 0.001).

Inclusiveness

Table 16: Regression 6.1 – High versus moderate inclusiveness logistic model

Variable	Standard error	Significance level	Odds ratio
Not exposed to *Sema Kenya*	Ref	-	-
Regularly exposed to *Sema Kenya*	.397	.016	2.611
Age (continuous)	.003	.023	1.007
Income – cannot afford food/clothes	Ref	-	-
Can afford food and clothes but purchasing of durables is difficult	.109	.523	1.072
Can afford main household appliances but purchasing a car is beyond our means	.128	.000	1.724
Can't afford house/no financial problems	.200	.001	1.905
Region - Rift Valley	Ref	-	-
Central	.140	.002	1.559
Coast	.168	.000	2.538
Eastern	.131	.000	1.694
Nairobi	.171	.000	2.452
North Eastern	.229	.000	2.370
Nyanza	.138	.914	1.015
Western	.157	.035	1.391
Interest in politics - not at all interested	Ref	-	-
Not very interested	.137	.518	1.093
Somewhat interested	.128	.970	1.005
Very interested	.127	.001	1.529
Unexposed – no food/clothes	Ref	-	-
Exposed – no durables	.475	.422	.683
Exposed – no cars	.483	.009	.283
Exposed – no house/no problems	.637	.524	.667
Constant	.183	.000	.493

The Nagelkerke R^2 statistic for this model was 0.078. The Hosmer and Lemeshow statistic had a chi-square of 2.766 and a significance level of 0.948. The Nagelkerke statistic indicates that the model only explains a limited proportion of the data. The Hosmer and Lemeshow statistic was not significant, so the model is unlikely to be mis-specified. Nonetheless, due to the low Nagelkerke R^2, results should be interpreted with caution.

3. Full results of structural equation modelling (SEM) analysis

This technical appendix summarises the results of the SEM analysis. See the full technical appendix for more information on the methodology and results, available at:
http://dataportal.bbcmediaaction.org/site/

Results from theoretical schematic 1
Figure 1 below gives the results from the first theoretical model, when modelled using the endline dataset. In all cases *** means significant at the 99.9% level, ** is significant at the 99% level and * is significant at the 95% level. No stars means outside the standard level of significance.

Figure 1: Model results from theoretical model 1
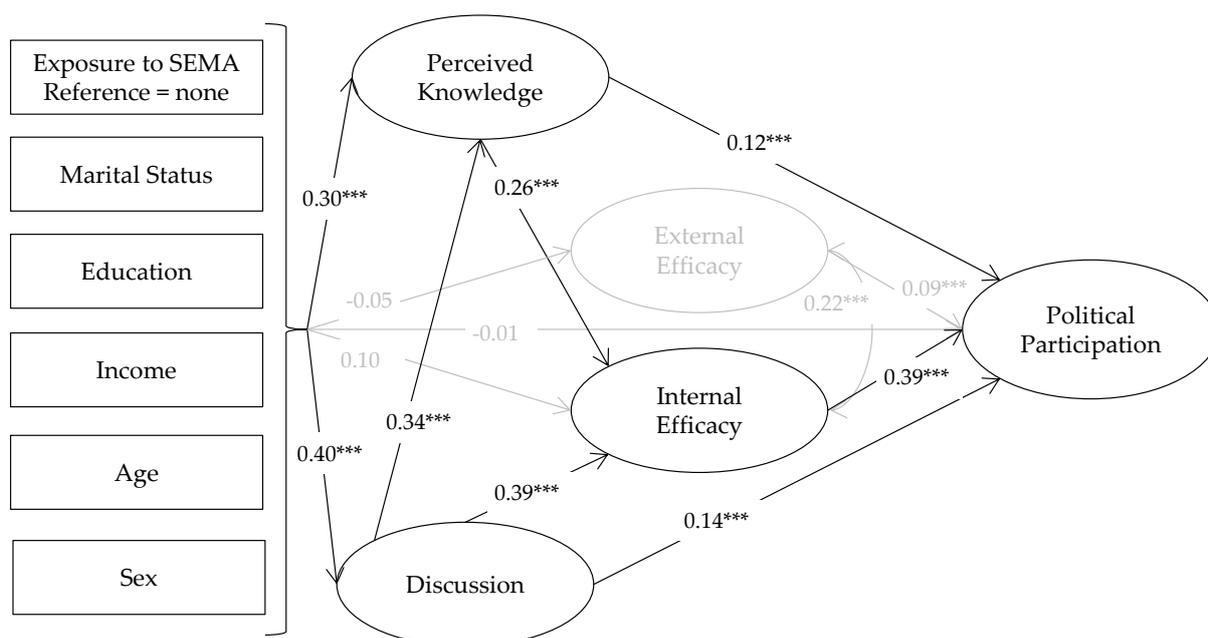
The coefficients presented on the diagram relate to exposure to *Sema Kenya* when compared to non-exposure. So the value of 0.30 from the control variables on the left hand side of the diagram to perceived knowledge means that those who were exposed to *Sema Kenya* reported that they had political knowledge about a third of standard deviation (0.30) higher than those who were not exposed. Similarly, levels of political discussion were higher by two-fifths (0.40) among those exposed to the programme (via either TV or radio).

The values between the ovals represent the standardised relationships between the latent variables, when controlling for all of the elements on the left hand side of the diagram. So the 0.26 between perceived knowledge and internal efficacy means that those who were one standard deviation higher on political knowledge were likely to be around a quarter of a standard deviation (0.26) higher on their feeling of political internal efficacy.

The components that are 'greyed out' were non-significant mediator variables and pathways between exposure to *Sema Kenya* and political participation. Put another way, these components played no role in shaping the political participation of the individual due to exposure to *Sema Kenya*.

There were direct and indirect effects of the control variable on the ultimate dependent variable (political participation). These were found by multiplying the coefficients along the pathways represented by single-headed arrows. The rationale for the pathways was given by the theoretical

models outlined in a separate report (available on request). Hence, the results are valid only in as much as the theoretical arguments are valid.

These results show that:
- The total effect of *Sema Kenya* on political participation = 0.239****
  - This means that exposure to *Sema Kenya* was associated with increased political participation by 0.24 of a standard deviation
- The direct effect of *Sema Kenya* on political participation = -.0.007 (not significant)
  - This means that the direct effect was slightly negative, but insignificant

Indirect effects, in order of size:
1. Via political discussion and then internal efficacy 0.061***
2. Via political discussion = 0.057***
3. Via knowledge = 0.034**
4. Via knowledge and then internal efficacy = 0.030***
5. Via knowledge and then political discussion = 0.016**
6. Via political discussion, knowledge and internal efficacy = 0.014***
7. All other pathways (via internal efficacy or external efficacy) not significant

Overall, therefore, exposure to *Sema Kenya* was associated with an increase in political participation but this was found to operate through the associated increases in the mediating variables in the model. The most relevant of these was through increases in political discussions and, to a lesser extent, through increased political knowledge. Internal political efficacy increased with both increases in discussion and knowledge, and this was associated with a further increase in political participation.

There were both direct and indirect effects from the other control variables included in the model. The total direct effects of each followed the below pattern:
- Being educated to a level lower than secondary school reduced individuals' political participation, while being educated to secondary school and above increased participation
- Being currently married was associated with a higher level of political participation, relative to those who were single
- The wealthiest people (i.e. those in the highest two income categories) participated most, all others participated to the same extent as those in the 'middle-to-low' income category (i.e. those reporting: "We can afford food and clothes, but purchasing of durables such as TV set or a refrigerator is difficult for us")
- Age (being older) increased participation
- Women participated less than men (by about 0.3 of a standard deviation)

Results from theoretical schematic 2

Figure 2 below gives the results from the second theoretical model. The presentation is of the same form as the previous section but some extra indirect pathways were included due to the addition of pathways from perceived political knowledge to external political efficacy, and from external political efficacy to internal political efficacy.

Figure 2: Model results from theoretical model 2
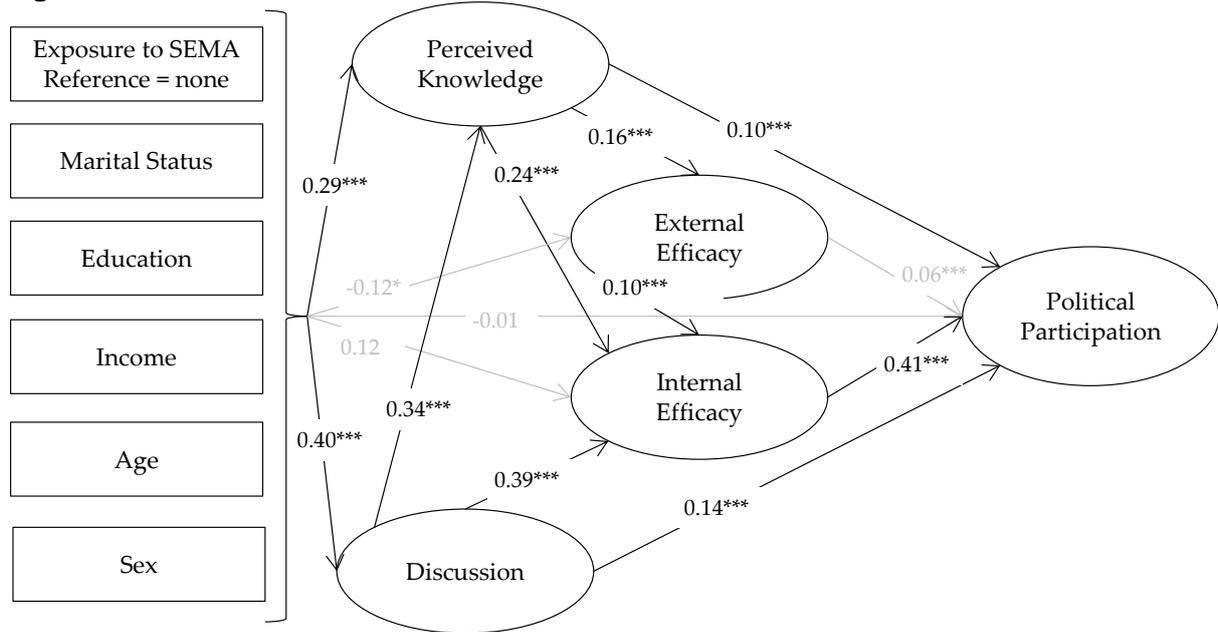
These results show that:
- The total effect of *Sema Kenya* on political participation = 0.239****
  - As in model 1, this means that exposure to *Sema Kenya* was associated with an increase in political participation of 0.24 of a standard deviation
- The direct effect of *Sema Kenya* on political participation = -.0.005 (not significant)
  - This again means that the direct effect was slightly negative, but insignificant

Indirect effects, in order of size:
1. Via political discussion and then internal efficacy = 0.063***
2. Via political discussion = 0.055***
3. Via knowledge = 0.029**
4. Via knowledge and then internal efficacy = 0.029***
5. Via knowledge and then political discussion = 0.014**
6. Via political discussion, knowledge and internal efficacy = 0.014***
7. Via knowledge and then external efficacy = 0.003*
8. Via knowledge, external efficacy then internal efficacy = 0.001*
9. All other pathways (via internal efficacy or external efficacy) not significant

These results are consistent with those found in the previous model and thus reiterate the same conclusions: that *Sema Kenya* only affected political participation through indirect pathways and that increases in discussion and, to a slightly lesser extent, knowledge, were most salient for this relationship.

All of the effects of the control variables were the same as in the first model.