Technical appendix

Strengthening accountability through media in Bangladesh: final evaluation

July 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 Bangladesh: final evaluation*. This is available from: http://dataportal.bbcmediaaction.org/site/assets/uploads/2016/09/Bangladesh-Country-Report-2017.pdf

1. Survey sampling methodology

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

Table 1: Quantitative surveys conducted (2012-2015)

Study	Data collection	Sample size	Criteria
Baseline survey	September–October 2012	n=5,628	Nationally-representative, adults 15+
Midline survey	April–May 2013	n=3,550	Nationally-representative, adults 15+
Reach survey 2013	June 2013	n=3,440	Nationally-representative, adults 15+
Reach survey 2014	June–July 2014	n=2,600	Nationally-representative, adults 15+
Endline survey	May–June 2015	n=2,650	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 latest census data. 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 vs rural), nested weights applied.

2. Regression analysis

This technical appendix summarises the results of analysis BBC Media Action carried out on the Bangladesh endline dataset (2015), which was representative of Bangladesh's national adult population (aged 15+).

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

Variables

The independent variable for regression analysis was exposure to *Sanglap*, with two categories: regularly exposed to *Sanglap* (exposed to at least every other episode), and never exposed to *Sanglap*. 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 and participation)
- 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
All analysis was carried out on the Bangladesh endline dataset (2015).

Model	Model performance			Association with exposure		
	Sample size	R square	Significance (OLS only)	Association (OR/ coefficient)	95% confidence interval/ standard error	Significance
Regression 1: Knowledge OLS model	2,650	0.219	<0.001	1.006	0.645-1.367	<0.001
Regression 2: Internal efficacy OLS model	2,075	0.044	<0.001	0.170	0.083-0.258	<0.001
Regression 3: Discussion OLS model	2,084	0.096	<0.001	0.127	0.065-0.189	<0.001
Regression 4: Participation OLS model	2,650	0.204	<0.001	0.410	0.208-0.611	<0.001
Regression 5: Participation OLS model with gender interaction	2,650	0.206	<0.001	Exposure for women: -0.541	-1.050 to -0.31	0.038

Full model results

Knowledge

Table 3: Regression 1 – Knowledge OLS model

Predictor	Unstandardized coefficients		Standardized coefficient	Significance	95.0% confidence interval for B	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sanglap*	Ref	-	-	-	-	-
Regularly exposed to *Sanglap*	1.006	.184	.116	1.006	.184	.116
Group membership – not an active member	Ref	-	-	-	-	-
Active member	.741	.448	.033	.741	.448	.033
Interest in politics – not at all interested	Ref	-	-	-	-	-
Not very interested	.157	.165	.026	.157	.165	.026
Somewhat interested	.258	.163	.047	.258	.163	.047
Very interested	1.123	.200	.153	1.123	.200	.153
Region - Dhaka	Ref	-	-	-	-	-
Barisal	.845	.237	.075	.845	.237	.075
Chittagong	-1.128	.151	-.171	-1.128	.151	-.171
Khulna	.664	.196	.074	.664	.196	.074
Rajshahi	.176	.177	.023	.176	.177	.023
Rangpur	.039	.196	.004	.039	.196	.004
Sylhet	-.722	.235	-.068	-.722	.235	-.068
Male	Ref	-	-	-	-	-
Female	-.196	.122	-.037	-.196	.122	-.037
Rural	Ref	-	-	-	-	-
Urban	.411	.129	.065	.411	.129	.065
Age 15-24	Ref	-	-	-	-	-
Age 25-34	.144	.168	.024	.144	.168	.024
Age 35-44	.328	.191	.050	.328	.191	.050
Age 45-54	.276	.221	.033	.276	.221	.033
Age 55-64	.176	.251	.018	.176	.251	.018
Age 65+	.480	.302	.038	.480	.302	.038
Education - illiterate	Ref	-	-	-	-	-
Literate	.176	.172	.023	.176	.172	.023
Completed primary	.779	.179	.102	.779	.179	.102
Completed secondary	1.078	.160	.178	1.078	.160	.178
Completed college/university	2.187	.205	.271	2.187	.205	.271
Income - medium	Ref	-	-	-	-	-
Low	.058	.149	.008	.058	.149	.008
High	1.462	.451	.065	1.462	.451	.065
Marital status – married,	Ref	-	-	-	-	-

living with spouse						
Single	-.044	.192	-.006	-.044	.192	-.006
Married, not living with spouse	.038	.646	.001	.038	.646	.001
Divorced/separated	-.872	.968	-.018	-.872	.968	-.018
Widowed	-.469	.324	-.030	-.469	.324	-.030
In a marriage where the husband has more than one wife	-1.735	1.925	-.018	-1.735	1.925	-.018
Constant	3.690	.259	-	.000	3.182	4.198

The model had an adjusted R square of 0.219. The Durbin-Watson value was 1.573. The F statistic was 20.519 (significance < 0.001).

Table 4: Regression 2 – Internal efficacy OLS model
Dependent variable: average internal efficacy (1 to 4)

Predictor	Unstandardized coefficients		Standardized coefficient	Significance	95.0% confidence interval for B	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sanglap*	Ref	-	-	-	-	-
Regularly exposed to *Sanglap*	0.17	0.045	0.087	0.00	0.083	0.258
Male	Ref	-	-	-	-	-
Female	-0.172	0.027	-0.14	0.00	-0.226	-0.119
Age 15-24	Ref	-	-	-	-	-
Age 25-34	0.031	0.036	0.023	0.39	-0.039	0.101
Age 35-44	0.07	0.041	0.045	0.09	-0.011	0.15
Age 45-54	0.036	0.049	0.019	0.46	-0.059	0.132
Age 55-64	0.082	0.056	0.036	0.14	-0.027	0.191
Age 65+	-0.09	0.061	-0.036	0.14	-0.209	0.029
Social class - lower	Ref	-	-	-	-	-
Upper	0.086	0.061	0.033	0.16	-0.034	0.205
Middle	0.033	0.034	0.024	0.33	-0.033	0.099
Media consumption – less than weekly	Ref	-	-	-	-	-
At least weekly	0.109	0.086	0.028	0.20	-0.06	0.279
Urban	Ref	-	-	-	-	-
Rural	-0.022	0.031	-0.016	0.47	-0.083	0.039
Education – no schooling	Ref	-	-	-	-	-
Some primary	0.03	0.043	0.018	0.48	-0.053	0.114
Completed primary	0.075	0.044	0.044	0.09	-0.012	0.161
Completed secondary	0.126	0.04	0.093	0.00	0.047	0.204
Completed college/university	0.106	0.051	0.061	0.04	0.006	0.205
Constant	2.589	0.097	-	0.00	2.399	2.779

The model had an adjusted R square of 0.044. The F statistic was 7.292 (significance < 0.001).

Discussion

Table 5: Regression 3 – Discussion OLS model
Dependent variable: average frequency of discussion (1 to 3)

Predictor	Unstandardized coefficients		Standardized coefficient	Significance	95.0% confidence interval for B	
	Beta	Standard error	Beta		Lower bound	Upper bound
Not exposed to *Sanglap*	Ref	-	-	-	-	-
Regularly exposed to *Sanglap*	0.127	0.032	0.089	0.00	0.065	0.189
Male	Ref	-	-	-	-	-
Female	-0.175	0.019	-0.195	0.00	-0.213	-0.137
Age 15-24	Ref	-	-	-	-	-
Age 25-34	0.046	0.025	0.046	0.07	-0.004	0.095
Age 35-44	0.015	0.029	0.013	0.61	-0.042	0.072
Age 45-54	0.027	0.034	0.019	0.44	-0.041	0.094
Age 55-64	0.01	0.039	0.006	0.79	-0.067	0.088
Age 65+	-0.059	0.043	-0.032	0.17	-0.142	0.025
Social class - lower	Ref	-	-	-	-	-
Upper	0.083	0.043	0.044	0.06	-0.002	0.168
Middle	-0.002	0.024	-0.002	0.93	-0.049	0.045
Media consumption – less than weekly	Ref	-	-	-	-	-
At least weekly	-0.015	0.058	-0.006	0.79	-0.13	0.099
Urban	Ref	-	-	-	-	-
Rural	0.03	0.022	0.03	0.17	-0.013	0.073
Education – no schooling	Ref	-	-	-	-	-
Some primary	0.009	0.03	0.007	0.78	-0.051	0.068
Completed primary	0.149	0.031	0.12	0.00	0.088	0.211
Completed secondary	0.134	0.028	0.136	0.00	0.078	0.19
Completed college/university	0.213	0.036	0.168	0.00	0.142	0.283
Constant	1.558	0.066		0.00	1.428	1.689

The model had an adjusted R square of 0.096. The F statistic was 15.810 (significance < 0.001).

Participation

Table 6: Regression 4 – Participation OLS model
Dependent variable: average political participation (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 *Sanglap*	Ref	-	-	-	-	-
Regularly exposed to *Sanglap*	.410	.103	.084	.000	.208	.611
Group membership – not an active member	Ref	-	-	-	-	-
Active member	1.000	.251	.080	.000	.508	1.491
Interest in politics – not at all interested	Ref	-	-	-	-	-
Not very interested	-.003	.093	-.001	.975	-.184	.179
Somewhat interested	.271	.091	.087	.003	.092	.450
Very interested	.638	.112	.155	.000	.418	.857
Region - Dhaka	Ref	-	-	-	-	-
Barisal	.726	.133	.115	.000	.466	.986
Chittagong	-.183	.085	-.049	.030	-.350	-.017
Khulna	.204	.109	.040	.062	-.010	.419
Rajshahi	.241	.099	.055	.015	.047	.435
Rangpur	.142	.109	.029	.195	-.073	.357
Sylhet	.160	.132	.027	.224	-.098	.418
Male	Ref	-	-	-	-	-
Female	-.786	.068	-.261	.000	-.920	-.652
Rural	Ref	-	-	-	-	-
Urban	-.184	.072	-.052	.010	-.325	-.043
Age 15-24	Ref	-	-	-	-	-
Age 25-34	.255	.094	.077	.007	.070	.439
Age 35-44	.391	.107	.105	.000	.182	.601
Age 45-54	.489	.124	.104	.000	.246	.731
Age 55-64	.510	.140	.092	.000	.235	.785
Age 65+	.725	.169	.102	.000	.393	1.056
Education - illiterate	Ref	-	-	-	-	-
Literate	.068	.096	.016	.482	-.121	.256
Completed primary	.169	.100	.039	.092	-.027	.365
Completed secondary	.318	.089	.093	.000	.143	.493
Completed college/university	.628	.115	.138	.000	.403	.853
Income - medium	Ref	-	-	-	-	-
Low	.140	.083	.035	.093	-.023	.304
High	.505	.252	.040	.045	.011	1.000
Marital status – married, living with spouse	Ref	-	-	-	-	-

Single	-.099	.107	-.025	.356	-.309	.111
Married, not living with spouse	.031	.361	.002	.931	-.677	.739
Divorced/separated	-.443	.542	-.016	.413	-1.505	.619
Widowed	-.028	.181	-.003	.876	-.384	.327
In a marriage where the husband has more than one wife	-.863	1.076	-.016	.423	-2.974	1.248
Constant	2.830	.145	-	.000	2.546	3.114

The model had an adjusted R square of 0.204. The Durbin-Watson value was 1.663. The F statistic was 19.443 (significance < 0.001).

Table 7: Regression 5 – Participation OLS model with gender interaction
Dependent variable: average political participation (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 *Sanglap*	Ref	-	-	-	-	-
Regularly exposed to *Sanglap*	.510	.114	.105	.000	.287	.733
Male	Ref	-	-	-	-	-
Female	-.753	.070	-.250	.000	-.890	-.616
Exposure for women	-.541	.260	-.046	.038	-1.050	-.031
Group membership – not an active member	Ref	-	-	-	-	-
Active member	.976	.251	.078	.000	.485	1.468
Interest in politics – not at all interested	Ref	-	-	-	-	-
Not very interested	.005	.093	.002	.954	-.176	.187
Somewhat interested	.274	.091	.088	.003	.095	.453
Very interested	.637	.112	.154	.000	.417	.856
Region - Dhaka	Ref	-	-	-	-	-
Barisal	.712	.133	.113	.000	.452	.972
Chittagong	-.191	.085	-.051	.024	-.357	-.025
Khulna	.193	.109	.038	.078	-.021	.408
Rajshahi	.224	.099	.051	.024	.030	.419
Rangpur	.129	.110	.026	.241	-.086	.343
Sylhet	.157	.132	.026	.232	-.101	.415
Rural	Ref	-	-	-	-	-
Urban	-.180	.072	-.051	.012	-.321	-.039
Age 15-24	Ref	-	-	-	-	-
Age 25-34	.251	.094	.076	.008	.066	.435
Age 35-44	.388	.107	.104	.000	.179	.597
Age 45-54	.487	.124	.104	.000	.244	.729
Age 55-64	.514	.140	.093	.000	.240	.789

Age 65+	.725	.169	.102	.000	.394	1.057
Education - illiterate	Ref	-	-	-	-	-
Literate	.064	.096	.015	.509	-.125	.252
Completed primary	.168	.100	.039	.093	-.028	.364
Completed secondary	.324	.089	.095	.000	.149	.499
Completed college/university	.624	.115	.137	.000	.399	.849
Income - medium	Ref	-	-	-	-	-
Low	.139	.083	.035	.097	-.025	.302
High	.522	.252	.041	.038	.028	1.016
Marital status – married, living with spouse	Ref	-	-	-	-	-
Single	-.104	.107	-.026	.330	-.314	.106
Married, not living with spouse	.015	.361	.001	.966	-.692	.723
Divorced/separated	-.457	.541	-.017	.399	-1.518	.604
Widowed	-.038	.181	-.004	.835	-.393	.318
In a marriage where the husband has more than one wife	-.862	1.075	-.016	.423	-2.971	1.247
Constant	2.818	.145	-	.000	2.534	3.102

The model had an adjusted R square of 0.206. The Durbin-Watson value was 1.663. The F statistic was 19.443 (significance < 0.001).