Appendix for: International Institutions and Contests Over Compliance

May 2, 2014

Abstract

This appendix is divided into two sections, theoretical and empirical. The theoretical appendix shows all proofs and derivations for the formal model. The empirical appendix describes data, tests, and robustness checks in greater detail. For readers interested in the robustness checks described in the main text, the flexible estimation analysis begins on page 14, the logit algorithm is on page 15, and the placebo tests are on page 19.

Theoretical Appendix

This section contains the proofs for the formal model. For simplicity, I first characterize optimal effort levels in any subgame perfect Nash equilibrium. As noted in the text, this proofing strategy follows Corchón (2007). I then include the institution and show existence of the equilibrium I discussed.

Optimal Effort Levels

First, redefine group i's optimization problem as follows, by dividing its payoffs by V_i:

$$max_{e_i} \Pi_i(e_i, e_j)$$

$$max_{e_i} \frac{e_i}{e_i+e_j} V_i - c_i * e_i$$

$$max_{e_i} \frac{e_i}{e_i+e_j} - \frac{c_i}{V_i} e_i$$

$$max_{e_i} \frac{e_i}{e_i+e_j} - d_i e_i$$

Differentiating with respect to e_i yields:

$$\frac{e_j}{(e_i+e_j)^2} = d_i$$

Note, summing the two groups' first order conditions and simplifying yields:

$$\frac{e_i+e_j}{(e_i+e_j)^2} = d_i + d_j$$

$$e_i + e_j = \frac{1}{d_i+d_j}$$

Using equation this summation and the first order condition yields e_i^* as a function of d_i and d_j and Proposition 2.

$$e_i^* = d_j(e_i + e_j)^2$$

(From the FOC)

$$e_i^* = d_j[(\frac{1}{d_i+d_j})^2]$$

(From the summation, substituting)

$$e_i^* = \frac{d_j}{(d_i+d_j)^2}$$

For Corollary 1, this expression generates comparative statics relating d_i (and by extension c_i and V_i) to the optimal effort level, e_i^*.

$$\frac{\partial e_i^*}{\partial d_i} = \frac{-2d_j}{(d_i+d_j)^3}$$

We can also generate comparative statics relating d_j to i's optimal effort level.

$$\frac{\partial e_i^*}{\partial d_j} = \frac{d_i - dj}{(d_i+d_j)^3}$$

Substituting the optimal effort levels into the contest success function and simplifying yields Proposition 3. Taking derivatives yields Corollary 2.

Optimal Effort Levels With/Without Institutional Signal

We can express the effects of an institutional signal (or absence of signal) by using the results above and incorporating the effect of the signal on the PC group's expected value of winning the contest. Recall, the PC group's prior expected value to winning is $V_{PC} = pv_{PC}$, and its "prior" d_{PC} is $d_{PC} = \frac{c_{PC}}{pv_{PC}}$

For ease of notation, let $\gamma' \equiv \frac{1+2pq-q-p}{pq}$ and $\gamma'' \equiv \frac{p+q-2pq}{(1-q)p}$. Using Bayes rule, the PC group's updated beliefs that compliance is beneficial, after a signal are:

$$Pr(B|S) = \frac{pq}{pq+(1-p)(1-q)}$$

Using this expression, we can write the PC group's "updated" d_{PC} as:

$$d'_{PC} = d_{PC}\frac{1+2pq-q-p}{pq}$$
$$d'_{PC} = d_{PC}\gamma'$$

Similarly, when no signal is sent, the pro-compliance group updates its beliefs and expected value, denoted d''_{PC}.

$$Pr(B| \sim S) = \frac{p(1-q)}{p(1-q)+(1-p)q}$$
$$d''_{PC} = d_{PC}\frac{p+q-2pq}{(1-q)p}$$
$$d''_{PC} = d_{PC}\gamma''$$

This allows us to simplify the optimal effort levels of the PC and AC groups, with and without the signal.

$$e'^{*}_{PC} = \frac{d_{AC}}{(\gamma' d_{PC}+d_{AC})^2}$$
$$e'^{*}_{AC} = \frac{\gamma' d_{PC}}{(\gamma' d_{PC}+d_{AC})^2}$$

$$e''^{*}_{PC} = \frac{d_{AC}}{(\gamma'' d_{PC}+d_{AC})^2}$$
$$e''^{*}_{AC} = \frac{\gamma'' d_{PC}}{(\gamma'' d_{PC}+d_{AC})^2}$$

Equilibrium Winning Probabilities and Institutional Utility

The equilibrium winning probabilities described in Proposition 3 and in Corollary 2 follow directly from the optimal effort levels described above and the contest success function. And this expression is general to any d, so it can be modified to account for institutional signals by adding the appropriate γ to the appropriate place.

$$\phi_i(e_i^*, e_j^*) = \frac{d_j}{d_i+d_j}$$

The institution's expected utility for sending a signal given that it gets a positive private signal is:

$$EU_I(S|b) = Pr(B|b)\phi_{PC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})V_I - \phi_{AC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})l - k$$

The institution's expected utility for not sending a signal given a positive private signal is:

$$EU_I(\sim S|b) = Pr(B|b)\phi_{PC}(e^{*\prime\prime}_{PC}, e^{*\prime\prime}_{AC})V_I$$

Combing these two expressions yields Proposition 4:

$$EU_I(S|b) - EU_I(\sim S|b) = Pr(B|b)[\phi_{PC}(e^{*\prime}_{PC}, e^{*\prime}_{AC}) - \phi_{PC}(e^{*\prime\prime}_{PC}, e^{*\prime\prime}_{AC})]V_I - \phi_{AC}(e^{*\prime}_{PC}, e^{*\prime}_{AC})l - k$$

Taking derivatives and simplifying (and noting that $Pr(B|b) = \frac{p}{\gamma'}$) yields Corollary 3:

$$\frac{\partial EU_I(S) - EU_I(\sim S)}{\partial d_{PC}} = \frac{pV_I}{\gamma'}\left[\frac{\gamma'' d_{AC}}{(\gamma'' d_{PC} + d_{AC})^2} - \frac{\gamma' d_{AC}}{(\gamma' d_{PC} + d_{AC})^2}\right] - l\left[\frac{\gamma'}{\gamma' d_{PC} + d_{AC}} - \frac{(\gamma')^2 d_{PC}}{(\gamma' d_{PC} + d_{AC})^2}\right]$$

Existence of Equilibrium

The conditions for the informative equilibrium to exist are described here. The institution's signal must induce a large enough change in the PC group's effort levels to justify the institution's fixed costs and the risk of legitimacy loss. The magnitude of the signal's effect on the PC group's efforts is a function of the accuracy of the institution's private information, q, the relative valuations and costs of winning for the two groups, and the groups' prior beliefs about the expected value of compliance. The institution's costs to sending a public signal (both k and l) must be high enough to deter the institution from wanting to send a positive public signal even when it does not receive a positive private signal. And they must be low enough so that the institution wants to send the signal when it receives a positive private signal. Similarly, the PC group's costs to effort must be high enough to keep them from exerting the high level of effort ($e^{*\prime}_{PC}$) even when the institution does not send a signal. And the costs must be low enough to make the PC group want to choose the higher level of effort when they do observe the institution's signal.

The conditions for the existence of the equilibrium in Proposition 1 are:

1. $q > \frac{1}{2}$
2. $Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I \geq \phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B|\sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$
3. $Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_{PC} \geq c_{PC}(e'_{PC} - e''_{PC}) \geq Pr(B|\sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_{PC}$

The second condition comes from the institution's decision (a) to send the signal when they receive a positive private signal:

$$EU_I(S|b) \geq EU_I(\sim S|b)$$
$$Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_I - \phi_{AC}(e'_{PC}, e'_{AC})l - k \geq Pr(B|b)\phi_{PC}(e''_{PC}, e''_{AC})V_I$$
$$\phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B| \sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

and (b) to not send the signal when they do not receive a positive private signal:

$$EU_I(\sim S| \sim b) \geq EU_I(S| \sim b)$$
$$Pr(B| \sim b)\phi_{PC}(e''_{PC}, e''_{AC})V_I \geq Pr(B| \sim b)\phi_{PC}(e'_{PC}, e'_{AC})V_I - \phi_{AC}(e'_{PC}, e'_{AC})l - k$$
$$\phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B| \sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

Combining conditions (a) and (b) yields:

$$Pr(B|b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I \geq \phi_{AC}(e'_{PC}, e'_{AC})l + k \geq Pr(B| \sim b)[\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC})]V_I$$

Note that $q > \frac{1}{2}$ ensures that $Pr(B|b) > Pr(B| \sim B)$. Also, Proposition 2 and 3 guarantee that $\phi_{PC}(e'_{PC}, e'_{AC}) - \phi_{PC}(e''_{PC}, e''_{AC}) > 0$. Since $V_I > 0$, there exist a pair $\{l, k\}$ small enough for both conditions (2) to be met.

The third condition is similar, but for the PC group. It says that the PC group wants to exert "high effort" iff they observe a positive institutional signal and low effort iff they do not observe this signal. The two analogous expressions are (a):

$$EU_{PC}(e'^{*}_{PC}|b) \geq EU_{PC}(e''^{*}_{PC}|b)$$
$$Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_{PC} - c_{PC}e'^{*}_{PC} \geq Pr(B|b)\phi_{PC}(e''_{PC}, e''_{AC})V_{PC} - c_{PC}e''^{*}_{PC}$$

and (b):

$$EU_{PC}(e''^{*}_{PC}| \sim b) \geq EU_{PC}(e'^{*}_{PC}| \sim b)$$
$$Pr(B| \sim b)\phi_{PC}(e''_{PC}, e''_{AC})V_{PC} - c_{PC}e''^{*}_{PC} \geq Pr(B|b)\phi_{PC}(e'_{PC}, e'_{AC})V_{PC} - c_{PC}e'^{*}_{PC}$$

Conditions (a) and (b) combine for condition (3) above. Meeting this condition requires that the costs of effort, relative to the value of winning the prize, be "just right." They have to be small enough to allow the PC group to increase its effort after a signal and large enough to keep it from simply exerting that high effort level regardless of the signal.

$$Pr(B|b)[\phi_{PC}(e'^*_{PC}, e'^*_{AC}) - \phi_{PC}(e''^*_{PC}, e''^*_{AC})]V_{PC} \geq c_{PC}(e'^*_{PC} - e''^*_{PC} \geq Pr(B| \sim b)[\phi_{PC}(e'^*_{PC}, e'^*_{AC}) - \phi_{PC}(e''^*_{PC}, e''^*_{AC})]V_{PC}$$

Empirical Appendix

The main text's description of the empirical tests is necessarily brief. Here I give a more detailed description of all the tests run and the results. I first describe the full details of the survey questions and procedure that produced the data. I then describe the procedure and results for the main probit regression tests, and then describe the procedure and results for the three robustness checks.

Survey Data

The data used here come from a set of nationally representative surveys conducted by Infotrak. Infotrak is a Kenyan polling firm which is associated with Harris Interactive Inc. According to Infotrak, the survey samples were each designed using Population Proportionate to Size sampling, using the 2009 Kenyan Population and Housing Census as the sample frame for national representativeness. Infotrak used the "district" as the administrative boundary for sampling. Surveys were conducted at the household level with face to face interviews. 25% of interviews were back-checked for quality control.[1]

The surveys that I analyzed were conducted on the following dates, with the number of individuals surveyed in parentheses: December 2010 (1,543); January 2011 (1,500); June 2011 (1,905); July 2011 (1,611); August 2011 (1,020); and October 2011 (1,477). The effective sample size

[1] Infotrak-Harris Popularity Poll Report Summaries.

used in the regressions varies slightly from these numbers since some respondents may not have responded to all questions and in the North Eastern region, for two surveys, no respondents selected Kenyatta.

The main question analyzed asked respondents "Apart from President Mwai Kibaki, who would you vote for as your President if presidential elections were held today?" Respondents chose from a list which included the top 10-13 candidates, as well as an option to specify "other" or "undecided." In some surveys, Infotrak asked a slightly different variation of this question: "If elections were held today, who would you vote for as president and as vice president?" and again used a show card with the top candidates to have respondents answer. I use this question to construct the dummy variable, k_i, which equals one if respondent i selected Kenyatta as their preferred candidate and zero otherwise. In the summary statistics table, this variable is labelled *Kenyatta*. I also coded analogous dummy variables for Odinga, o_i, which are reported in the summary statistics table under *Odinga*. The Odinga variables are used in some of the robustness checks.

The survey also asked a set of demographic questions that I use in the analysis. *Male* is a dummy variable that equals one if the respondent was male. Respondents were also classified by whether they lived in urban or rural districts. *Urban* is a dummy variable that equals 1 for respondents who lived in urban areas. Respondents were asked whether their religion was Catholic, Protestant, Muslim, Hindu or Other. *Catholic* and *Protestant* are dummy variables indicating that the respondent chose Catholic or Protestant, respectively. Only about 12% of respondents were not Catholic or Protestant, so I left the other three categories together as the base category. *Age* is a categorical variable that classifies respondents by their age bracket. Some surveys used a 5 pt scale, while others used an 8 pt scale. I conformed the 8pt scale to the 5pt scale as closely as possible, though there are some mismatches. For example, on the 8pt scale, the first two age brackets are 18-20 and 21-25. On the 5pt scale, the first age bracket is 18-24. I re-coded the 8pt scale so that values of 1 or 2, referring to ages 18-25, were equal to 1 on the 5pt scale.[2]

[2]This "miscodes" 25 year old respondents. However, this is highly unlikely to affect results. Age is not a significant

Each respondent was also classified by their region of residence. The 8 regions are: Nairobi, Nyanza, Central, Rift Valley, North Eastern, Eastern, Coast, and Western. For each respondent, I let r_i^j be binary variable that equals 1 if respondent i lives in region j, and zero otherwise. *Monthnum* is a counter variable indicating the month that the survey was conducted in.

Table 1 shows summary statistics of the variables for all the surveys and then for each survey individually. The bottom part of the survey reports the number of respondents by survey and by region.

Probit Tests Algorithm

The basic approach of this algorithm is to use the following steps: (1) Use the data concerning respondents from the pre-summonses surveys (December 2010 and January 2011) to estimate a probit regression of support for Kenyatta on the observables that we know about the respondents like their demographic characteristics and region of residence. (2) Collect those coefficients and use them to predict that individual's latent support for Kenyatta and likelihood of supporting him, from the post-summonses data, based on his or her characteristics. (3) Calculate the difference difference between that individual's predicted support and whether they actually indicated support for Kenyatta. (4) Plot and analyze the relationship between predicted support and the differences.

Let X denote the matrix which contains all of the individuals' observable characteristics, with each row corresponding to one individual. Let X^{pre} denote the matrix only containing those individuals surveyed before the ICC summonses and X^{post} denote the matrix containing only those individuals surveyed after the ICC summonses. Let k denote the vector where the ith entry equals 1 if individual i supported Kenyatta. Let k^{pre} denote the vector containing only individuals from the pre-summonses surveys, and k^{post} denote individuals from the post-summonses surveys.

For step 1, as described in the text, I estimated a probit regression with the following specifica-

predictor of Kenyatta support. I have also re-estimated all models simply leaving the age values as they were in the original surveys. The pairwise correlation coefficient between my recoding and this approach is approximately 0.90. Results from analysis are not different in any meaningful way.

Table 1: Summary statistics

	Full Sample		**Dec 10**		**Jan 11**		**Jun 11**		**Jul 11**		**Aug 11**		**Oct 11**	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Kenyatta	0.149	0.356	0.041	0.199	0.133	0.34	0.186	0.389	0.172	0.377	0.157	0.364	0.183	0.387
Ruto	0.076	0.264	0.062	0.241	0.036	0.186	0.071	0.258	0.138	0.345	0.084	0.277	0.07	0.255
Odinga	0.431	0.495	0.367	0.482	0.462	0.499	0.467	0.499	0.458	0.498	0.378	0.485	0.414	0.493
Male	0.583	0.493	0.577	0.494	0.552	0.497	0.61	0.488	0.574	0.495	0.602	0.49	0.583	0.493
Age	2.022	1.004	2.303	1.081	1.607	0.807	1.829	0.902	1.954	0.981	2.185	0.993	2.439	1.038
Urban	0.592	0.491	0.52	0.5	0.602	0.49	0.652	0.477	0.561	0.496	0.626	0.484	0.570	0.495
Catholic	0.422	0.494	0.451	0.498	0.444	0.497	0.456	0.498	0.372	0.484	0.42	0.494	0.379	0.485
Protestant	0.458	0.498	0.404	0.491	0.435	0.496	0.433	0.496	0.485	0.5	0.473	0.5	0.523	0.5
Nairobi	863		108		157		199		158		105		136	
Coast	571		86		121		116		98		50		100	
N. Eastern	296		48		33		84		75		26		30	
Eastern	1,240		134		241		254		129		288		194	
Central	1,045		180		233		195		127		112		198	
R. Valley	1,825		229		293		477		399		105		322	
Western	731		110		160		128		69		124		140	
Nyanza	1,272		220		207		282		214		143		206	
N	7,843		1,115		1,445		1,735		1,269		953		1,326	

tion, using the individuals in X^{pre} and k^{pre}:

$$k_i^* = X_i\beta + \sum_{j=1}^{6} \gamma_j r_i^j + \sum_{j=1}^{6} (\delta_j r_i^j * t) + \epsilon_i \tag{1}$$

$$\epsilon_i \sim i.i.d.N(0,1) \tag{2}$$

$$k_i = \begin{cases} 1 & \text{if } k_i^* > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Table 2 describes the resulting estimates. Let $\hat{\beta}$ denote the vector containing these coefficients. For step 2, I post-multiplied the vector of coefficients from step 1, $\hat{\beta}$ by the matrix of observables for individuals in the post-summonses surveys, X^{post}. For each element in the vector, I calculated the normal CDF of that value. Denote the resulting vector as $\Phi(\hat{k})$, where the ith entry corresponds to the predicted probability that individual i would support Kenyatta. For step 3, I calculated the difference described in the text, $d_i = \Phi(\hat{k}_i) - k_i$.

Alternate Specifications for the Probit Tests

In the main text, I presented the loess smoothed differences using all post-summonses data and then using only the June 2011 survey data, excluding Nairobi. For the sake of thoroughness, I want to show that those two figures are not representing idiosyncratic trends. Figure 1, Figure 2, and Figure 3 show the equivalent of Figure 5 (from the main text), using all regions, “zooming in” by using only the 3 post-summonses surveys (June 2011, July 2011, August 2011), 2 post-summonses surveys (June 2011, July 2011), and then the 1 post-summonses survey (June 2011).

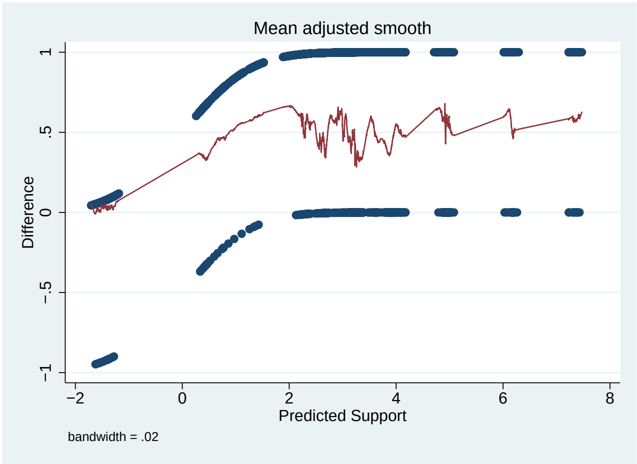
The trends shown in Figures 5 and 6 (main text) are also not artifacts of the loess smoothing bandwidth I chose. Figure 4 and Figure 5 show those same Figures from the main text with different smoothing bandwidths. The original bandwidths for Figure 5 was 0.1 and for Figure 6 was 0.03. The figures here show two higher and two lower bandwidths.

Table 2: Estimates from Probit Model, Pre-ICC Summonses Surveys

Male	.018 (.076)
Age	-.032 (.043)
Urban	-.075 (.086)
Catholic	.065 (.156)
Protestant	.141 (.154)
Nairobi	-13.312 (6.311)**
Coast	-1.828 (5.075)
Eastern	-8.824 (5.066)*
Central	-5.309 (4.285)
R. Valley	-5.893 (4.356)
Western	-6.306 (6.476)
Monthnum*Nairobi	1.212 (.392)***
Monthnum*Coast	.323 (.268)
Monthnum*Eastern	.864 (.263)***
Monthnum*Central	.638 (.155)***
Monthnum*R. Valley	.651 (.169)***
Monthnum*Western	.625 (.409)
Monthnum*Nyanza	.112 (.304)
N	2,479

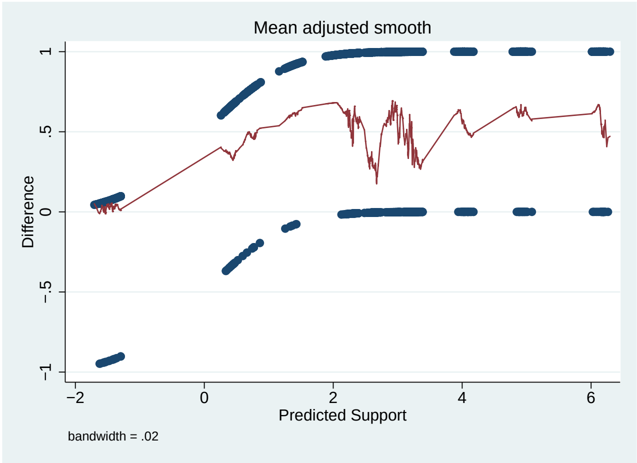
***: significance at 0.01 level; **: 0.05 level; *: 0.10 level.

Figure 1: Predicted versus Actual Support, First Three Post-Event Surveys Only, All Regions



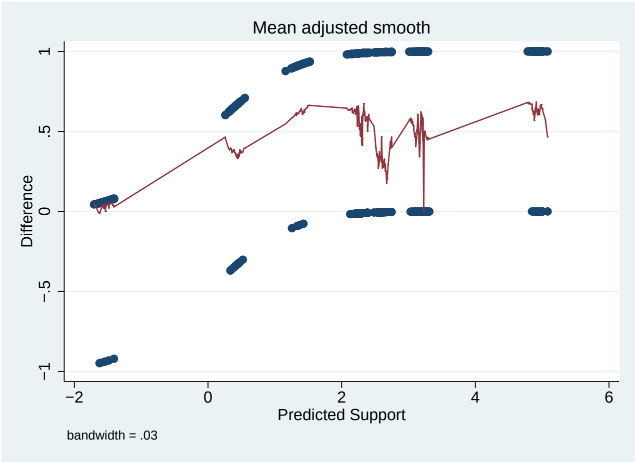
This figure only uses data from the June, July, August 2011 surveys and includes all regions. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 2: Predicted versus Actual Support, First Two Post-Event Surveys Only, All Regions



This figure only uses data from the June, July 2011 surveys and includes all regions. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 3: Predicted versus Actual Support, First One Post-Event Surveys Only, All Regions



This figure only uses data from the June 2011 survey and includes all regions. The horizontal axis is the linear prediction of latent support for Kenyatta using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 4: Predicted versus Actual Support, Alternate Bandwidths for Figure 5

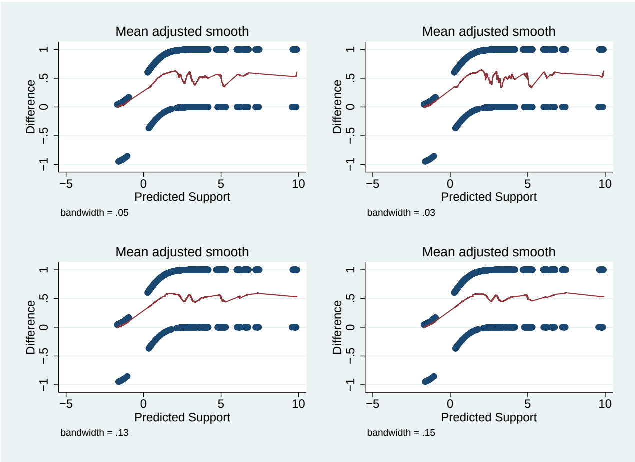
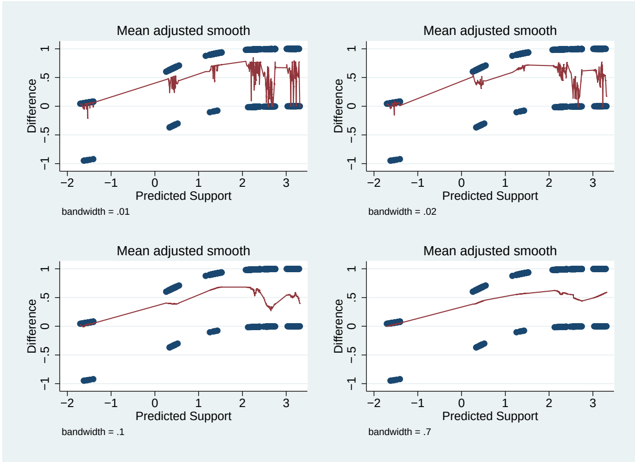


Figure 5: Predicted versus Actual Support, Alternate Bandwidths for Figure 6



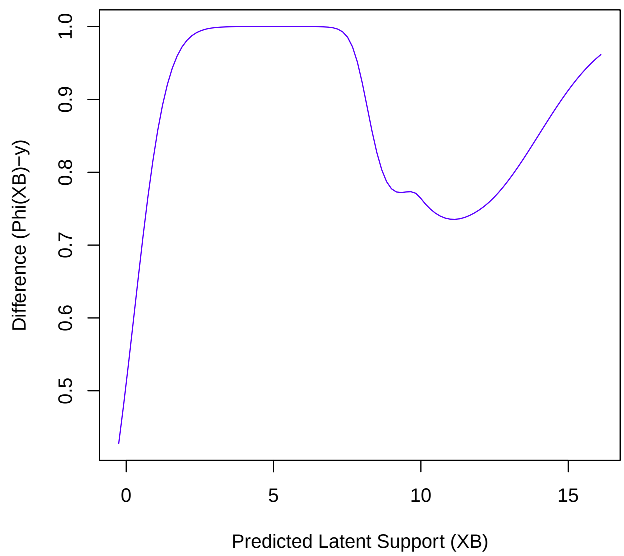
Flexible Estimation

The results are robust to flexible estimation techniques. The first step in the probit algorithm described above was to use data from before the ICC summonses to train a model of support for Kenyatta. For this step, I chose a particular functional form for the relationship between respondents' characteristics and their probability of supporting Kenyatta. While the functional form I chose was grounded in theoretical knowledge about this situation, it is important to establish that the results are not artifacts of my choices.

Kenkel and Signorino (2013) develop a technique in which the functional form for the effect of covariates on the outcome of interest is *estimated* rather than imposed. This approach incorporates the covariates, their polynomial expansions, and interactions the various terms them into a basis regression. It then selects the appropriate variables, expansions, and/or interactions using penalized regression. I apply their approach here by estimating the pre-summonses training model using their procedure via the *polywog* command in R. I then reconstruct the differences, d_i, as before, using the *polywog* estimates.[3] Figure 6 shows the a smoothed local fit line of those differences. The pattern is similar to the above results. The effect of the ICC is strongest in the middle and weaker

[3]Estimates were iterated and bootstrapped 500 times. All other specifications were left at the command's defaults.

Figure 6: Smoothed Predicted versus Actual Support, Flexible Estimation



This figure uses the *polywog* package (Kenkel and Signorino, 2013) to construct pre-ICC estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice. This figure shows the local fit loess line of those estimates.

on the leftmost and rightmost areas.

Since the main text did not report the coefficients that were returned by flexible estimate, I report them here, in Table 3

Logit Algorithm

Another of the robustness checks in the main text was to ensure that the distribution chosen for the disturbances in the probit tests wasn't influencing results. I repeated a similar algorithm, only using logit regressions and accompanying predicted probabilities. This section gives greater details on that procedure.

Table 3: Estimates from *polywog* on Pre-Summonses Data

	Coefficient
Intercept	-15.08382017
Nairobi	-4.92994391
Coast	0.91482249
Eastern	-9.62756830
Age	-0.15354126
Urban	-0.48163251
Month	0.88337292
Nairobi*Male	-0.07422260
Nairobi*Month	0.52040897
Coast*Urban	-0.43233870
Coast*Catholic	-0.42798632
Coast*Month	0.09906719
Eastern*Age	0.58829275
Eastern*Urban	-0.24688006
Eastern*Catholic	-0.82159971
Eastern*Protestant	-1.18679154
Eastern*Month	0.89299138
Central*Catholic	0.46716240
Central*Protestant	0.86386292
Central*Month	0.17765069
R. Valley*Protestant	0.26828779
R. Valley*Month	0.15369017
Western*Month	0.04860257
Male*Urban	0.45157589
Male*Month	-0.01684096
Urban*Protestant	0.49115883
Catholic*Month	0.02302705
N	2,586

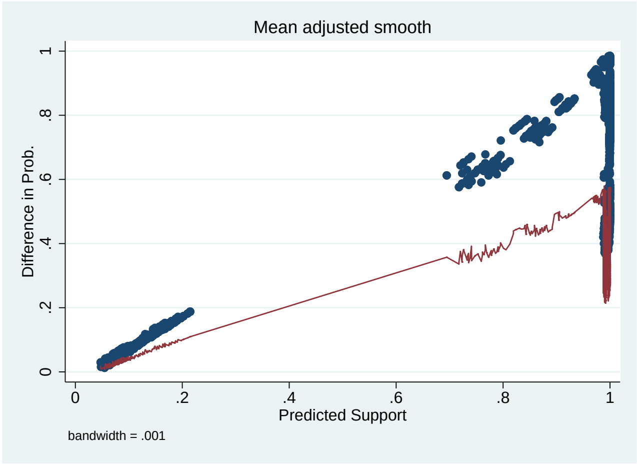
0.0.1 Logit Algorithm in Brief

As a second robustness check, I altered the algorithm above to remove any bias resulting from parametric assumptions I made about the distribution of individual level disturbances. In the algorithm above, I modeled an individual's likelihood of supporting Kenyatta as a function of that individual's latent support for Kenyatta and an individual level disturbance. Since I used a probit model, the disturbances were assumed to be normally distributed. The shape of the cumulative normal distribution might bias results in favor of finding the greatest effect of the summonses "in the middle," i.e. where the normal CDF is steepest. I want to check that the results are not artifacts of these "floor and ceiling effects," resulting from a parametric assumption I made.

To address this, I first used a logit regression to estimate a pre-summonses training model. I then estimated an analogous logit regression using individuals from the post-summonses surveys. For each post-summonses individual, I then calculated their predicted probability of supporting Kenyatta based on the estimates from the pre-summonses model and based on the estimates from the post-ICC model. I then calculated the difference: the predicted probability that an individual supported Kenyatta based on pre-summonses coefficients (from step 1) minus their probability based on post-summonses coeffcients (from step 2). This difference is interpreted in the same way as the differences in the probit approach above. A positive quantity shows that the individual's support for Kenyatta is lower, based on the post-summonses coefficients, than what we would have expected, based on the pre-summonses coefficients.

Figure 7 plots the results, with an individual's predicted probability of support based on pre-summonses coefficients on the horizontal axis and the difference in predicted probabilities on the vertical axis. The results are less smooth because of the large number of predicted probabilities clustered at 1. But crucially, the same pattern from above obtains. The effect of the ICC summonses is non-monotonically related to pre-summonses support in the way the theory predicts.

Figure 7: Predicted versus Actual Support, Logit Approach



The horizontal axis is the predicted probability of support for Kenyatta using pre-ICC logit estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta based on pre-ICC coefficients minus the predicted probability based on post-ICC coefficients. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

0.0.2 More Logit Algorithm Details

First, I estimated a logit regression of Kenyatta support on respondent observables, using the pre-summonses surveys. Table 4, Model 1, shows the resulting coefficients. Second, I estimated an analogous logit regression of Kenyatta support on respondent observables, using the post-summonses surveys. Table 4, Model 2, shows those resulting coefficients. Third, I constructed a similar difference to the one used in the probit tests. Specifically, for each respondent, I calculated their predicted probability of supporting Kenyatta based on the pre-summonses model, $p^{\hat{pre}} = \frac{1}{1+e^{-x_i \beta^{\hat{pre}}}}$. And I also calculated the same quantity based on the post-summonses model, $p^{\hat{post}} = \frac{1}{1+e^{-x_i \beta^{\hat{post}}}}$. I then calculated the difference in these predicted probabilities: $d_i = p^{\hat{pre}} - p^{\hat{post}}$. As with the probit tests, positive quantities correspond to lowered levels of support for Kenyatta than we would have expected, based on the pre-summonses training model.

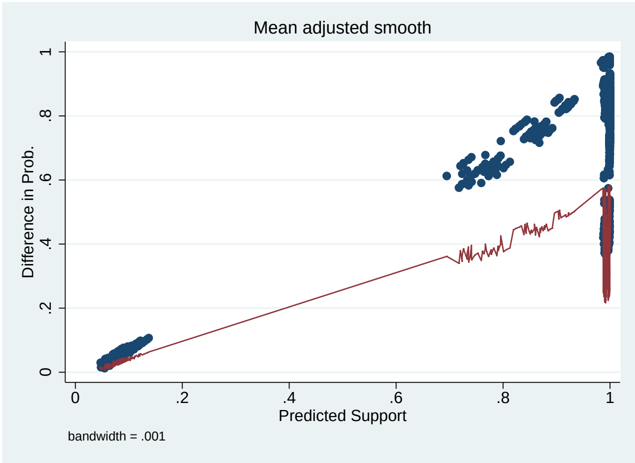
In the main text, I presented only the logit test results using data from all four of the post-summonses surveys. Figure 8, Figure 9, and Figure 10 "zoom-in," as above, using only the first three, first two, and first one post-summonses surveys.

Table 4: Estimates from Logit Model, Pre- and Post-ICC Summonses Surveys

	Pre	Post
	(1)	(2)
Male	.029	-.043
	(.143)	(.080)
Age	-.076	.060
	(.083)	(.040)
Urban	-.140	-.267
	(.164)	(.086)***
Catholic	.139	.569
	(.301)	(.176)***
Protestant	.306	.596
	(.297)	(.173)***
Nairobi	-29.863	-2.274
	(16.506)*	(3.005)
Coast	-2.835	4.308
	(11.910)	(3.593)
N. Eastern		2.363
		(3.713)
Eastern	-18.241	6.637
	(12.524)	(2.902)**
Central	-7.944	4.411
	(10.358)	(2.810)
R. Valley	-10.737	-1.401
	(10.645)	(2.779)
Western	-15.663	-1.725
	(16.917)	(4.454)
Nairobi*Month	2.731	.153
	(1.034)***	(.073)**
Coast*Month	.643	-.211
	(.551)	(.127)*
N. Eastern*Month		-.049
		(.137)
Eastern*Month	1.832	-.287
	(.618)***	(.063)***
Central*Month	1.118	-.102
	(.285)***	(.050)**
R. Valley*Month	1.278	.111
	(.345)***	(.046)**
Western*Month	1.541	.026
	(1.077)	(.179)
Nyanza*Month	.270	-.061
	(.772)	(.135)
N	2,479	5,283

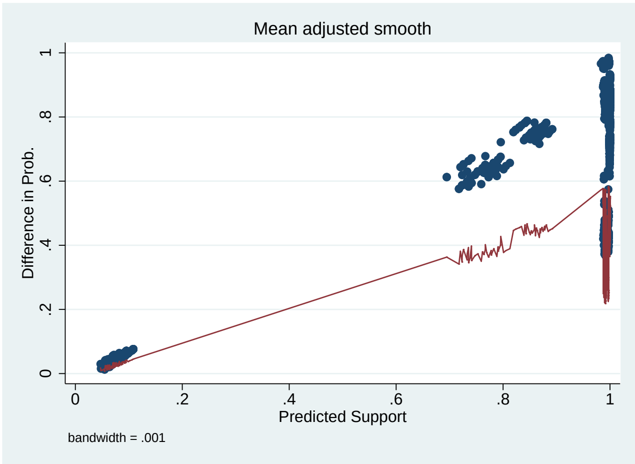
***: significance at 0.01 level; **: 0.05 level; *: 0.10 level.

Figure 8: Predicted versus Actual Support, Logit Approach, First Three Post-ICC Surveys



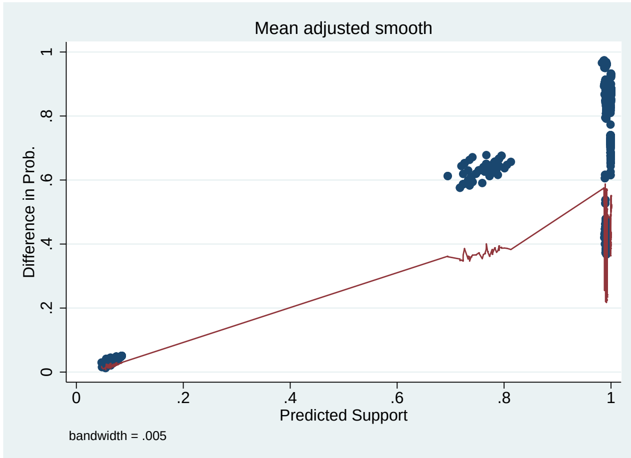
The horizontal axis is the predicted probability of support for Kenyatta using pre-ICC logit estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta based on pre-ICC coefficients minus the predicted probability based on post-ICC coefficients. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 9: Predicted versus Actual Support, Logit Approach, First Two Post-ICC Surveys



The horizontal axis is the predicted probability of support for Kenyatta using pre-ICC logit estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta based on pre-ICC coefficients minus the predicted probability based on post-ICC coefficients. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 10: Predicted versus Actual Support, Logit Approach



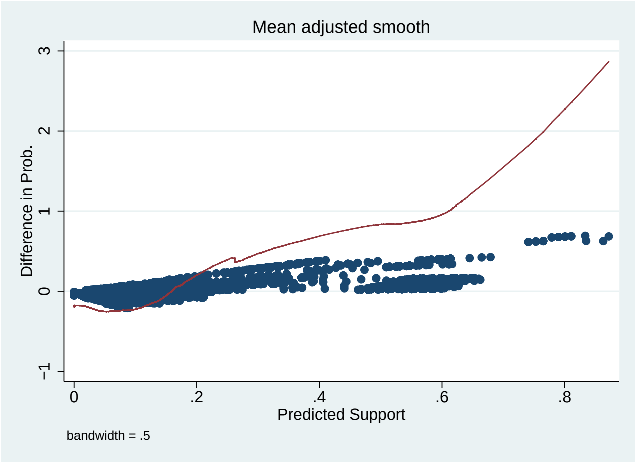
The horizontal axis is the predicted probability of support for Kenyatta using pre-ICC logit estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta based on pre-ICC coefficients minus the predicted probability based on post-ICC coefficients. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Placebo Test

As a third robustness check, I conducted a "placebo test" to further establish that the results are not mere artifacts of the approach I used. To do this, I make use of the fact that there are 4 post-ICC surveys. Specifically, I repeated the logit algorithm just described, but, instead of comparing pre- and post-summonses data, I used the first two post-summonses surveys and compared them to the last two post-summonses surveys. This is a placebo test, because, to the best of my knowledge, no event like the ICC's summonses occurred between the July 2011 and August 2011 surveys, i.e. between the first two post-summonses surveys and the latter two post-summonses surveys. So I should *not* expect to find the same relationship as above. This is indeed the case, as shown in Figure 11, where there is no strong pattern relating predicted support and ICC effects.

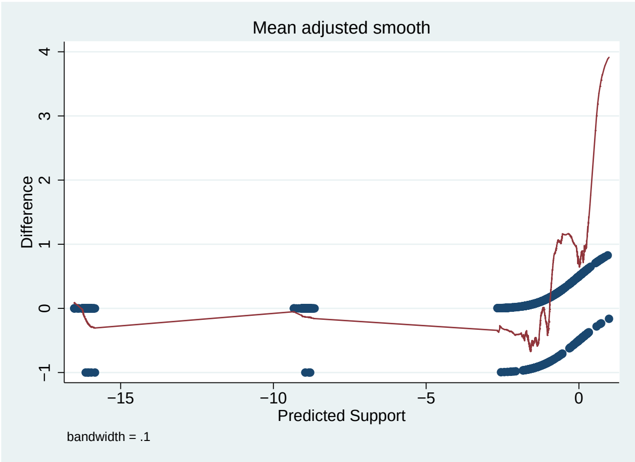
The above text presented the placebo test using the logit robustness algorithm. Figure 12 presents the same placebo test, only using the probit algorithm instead.

Figure 11: Predicted versus Actual Support, Logit Approach, Placebo Test



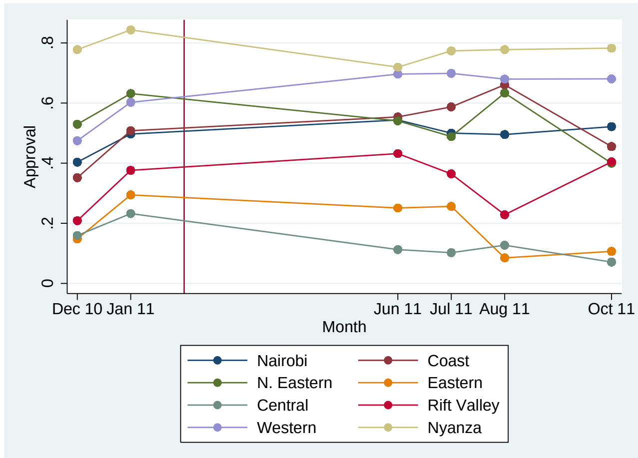
The horizontal axis is the predicted probability of support for Kenyatta using logit estimates from the first two post-ICC surveys. The vertical axis is the individual's predicted probability of supporting Kenyatta based on those coefficients minus the predicted probability based on coefficients from the second two post-ICC surveys. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 12: Predicted versus Actual Support, Probit Approach (Placebo)



The horizontal axis is the linear prediction of latent support for Kenyatta using the June, July 2011 estimates. The vertical axis is the individual's predicted probability of supporting Kenyatta minus the individual's observed choice, using the August, October 2011 data. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 13: Odinga Support by Region Over Time



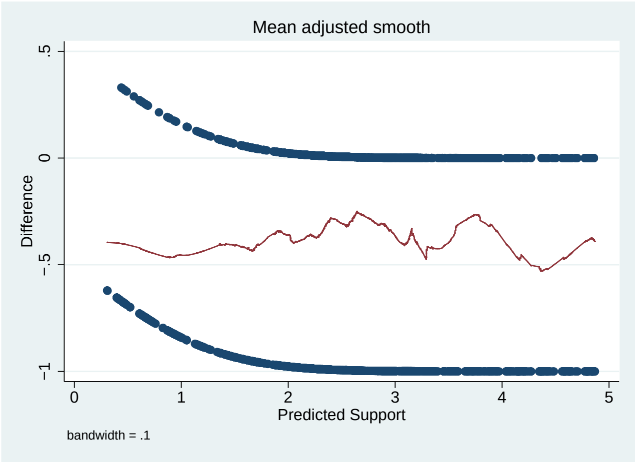
Percent of respondents in each survey answering that Odinga was their most preferred candidate for six surveys. Surveys were conducted by Infotrak in Kenya.

Odinga Tests

The main text focused on unexpected dips in support for Kenyatta which could possibly be attributed to the ICC's summonses. We can also detect similar unexpected increases in support for Kenyatta's main opponent, Raila Odinga. Figure 13 shows support for Odinga, by region, over time (similar to Figure 4 in the main text). I conducted the same analysis of differences for Odinga as I did for Kenyatta. There were only two changes. First, I used the dummy variable that indicated support for Odinga, o_i. Second, I constructed the differences so that positive values would represent unexpected increases in his support, $d_i = o_i - \Phi(\hat{o}_i)$. Positive quantities indicate that support for Odinga is higher than would be predicted based on estimates from the pre-ICC summonses data. This is to keep interpretation consistent with the Kenyatta analysis.

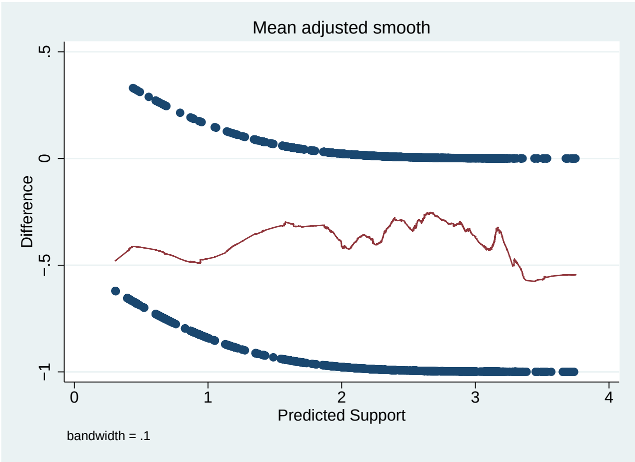
Figure 14 shows the Odinga version of Figure 5 from the main text. As predicted by the theory, and similar to the results for Kenyatta, the highest effect of the ICC appears to be in the middle portion of the plot. Figure 15, Figure 16, and Figure 17 "zoom in" on the ICC event as above, by using the first three, two, and one post-ICC surveys only. The same pattern is apparent in Figure 15 and Figure 16, although it is not as clear in Figure 17.

Figure 14: Predicted versus Actual Support (Odinga), All Post-Event Surveys



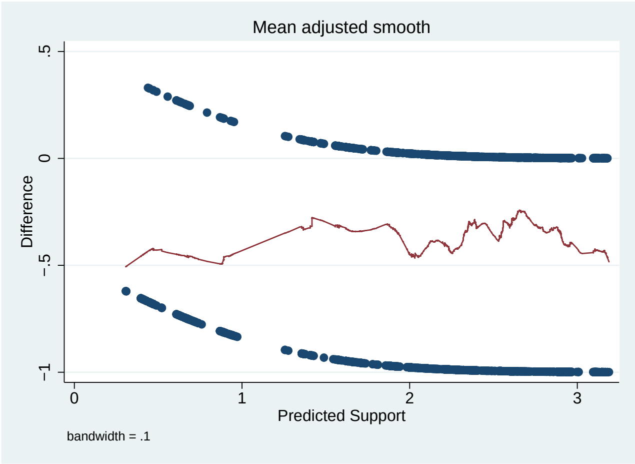
The horizontal axis is the linear prediction of latent support for Odinga using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Odinga minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 15: Predicted versus Actual Support (Odinga), First 3 Post-Event Surveys



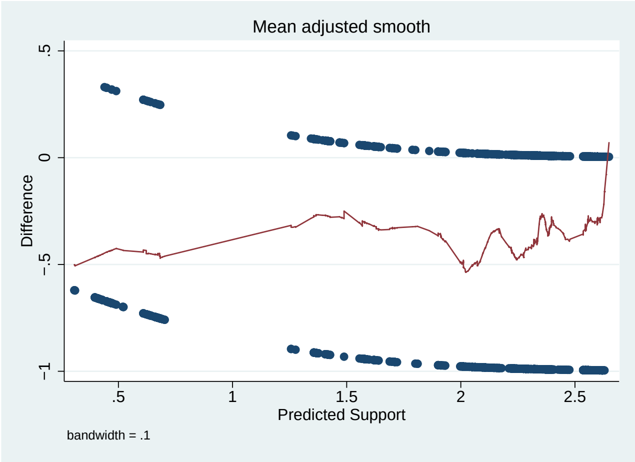
The horizontal axis is the linear prediction of latent support for Odinga using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Odinga minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 16: Predicted versus Actual Support (Odinga), First 2 Post-Event Surveys



The horizontal axis is the linear prediction of latent support for Odinga using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Odinga minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

Figure 17: Predicted versus Actual Support (Odinga), First 1 Post-Event Surveys



The horizontal axis is the linear prediction of latent support for Odinga using pre-ICC event estimates. The vertical axis is the individual's predicted probability of supporting Odinga minus the individual's observed choice. Smoothed loess line is included, where the mean of the smoothed values is constrained to equal the mean of the values on the vertical axis.

References

Corchón, Luis C. 2007. "The theory of contests: a survey." *Review of Economic Design* 11(2):69–100.

Kenkel, Brenton and Curtis S. Signorino. 2013. "Bootstrapped Basis Regression with Variable Selection: A New Method for Flexible Functional Form Estimation." Manuscript, University of Rochester.