**Supplemental Materials**

**A** **Reviewing Outcome Measurement in Literature on Voter Behavior**

To identify measurement patterns in the existing literature on voter behavior, we conducted a review of the top economic and political science journals. As search terms, we used “voter turnout” and “vote choice”, sometimes in combination with “experiment.” We report on this literature review in Table A.1 where we identify which studies use survey or administrative data (or both) to measure vote choice and voter turnout. We also identify whether the study was experimental in nature.

Table A.1: Literature on Vote Choice and Voter Turnout, by Measurement Type

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Author(s)  |  Year  |  Source  |  Experimental?  |  Measure of Vote Choice  |  Measure of Turnout  |
|  |  |  |  |  Survey  |  Administrative  | Survey  |  Administrative  |
| Giné & Mansuri  |  2018  |  American Economic Journal: Applied Economics  |  |  |  |  |  |
| Gianmarco  |  2017  |  Journal of Development Economics  |  |  |  |  |  |
| Mvukiyehe & Samii  |  2017  |  World Development  |  |  |  |  |  |
| Valenzuela & Michelson  |  2016  |  American Political Science Review  |  |  |  |  |  |
| Conroy-Krutz et al  |  2016  |  Comparative Political Science  |  |  |  |  |  |
| Nathan  |  2016  |  Comparative Political Science  |  |  |  |  |  |
| Bailey et al  |  2016  |  Political Behavior  |  |  |  |  |  |
| Shineman  |  2016  |  British Journal of Political Science  |  |  |  |  |  |
| Barton et al  |  2014  |  The Economic Journal  |  |  |  |  |  |
| Vicente  |  2014  |  The Economic Journal  |  |  |  |  |  |
| Chong et al  |  2014  |  The Journal of Politics  |  |  |  |  |  |
| Doherty & Adler  |  2014  |  Political Research Quarterly  |  |  |  |  |  |
| Fujiwara & Wantchekon  |  2013  |  American Economic Journal: Applied Economics  |  |  |  |  |  |
| Gerber et al  |  2013  |  American Journal of Political Science  |  |  |  |  |  |
| De La O  |  2013  |  American Journal of Political Science  |  |  |  |  |  |
| Utych & Kam  |  2013  |  The Journal of Politics  |  |  |  |  |  |
| Greene  |  2011  |  American Journal of Political Science  |  |  |  |  |  |
| Sheafer et al  |  2011  |  Comparative Political Science  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Author(s)  |  Year  |  Source  |  Experimental?  |  Measure of Vote Choice  |  Measure of Turnout  |
|  |  |  |  |  Survey  |  Administrative  |  Survey  | Administrative  |
| Panagopoulos  |  2011  |  The Journal of Politics  |  |  |  |  |  |
| Gerber et al  |  2010  |  American Political Science Review  |  |  |  |  |  |
| Arceneaux & Nickerson  |  2010  |  American Politics Research  |  |  |  |  |  |
| Arceneaux & Kolodny  |  2009  |  American Journal of Political Science  |  |  |  |  |  |
| Dale & Strauss  |  2009  |  American Journal of Political Science  |  |  |  |  |  |
| Gerber et al  |  2009  |  American Economic Journal: Applied Economics  |  |  |  |  |  |
| Bargsted & Kedar  |  2009  |  American Journal of Political Science  |  |  |  |  |  |
| Gerber et al  |  2008  |  American Political Science Review  |  |  |  |  |  |
| De La O & Rodden  |  2008  |  Comparative Political Science  |  |  |  |  |  |
| Krasno & Green  |  2008  |  The Journal of Politics  |  |  |  |  |  |
| Hellwig  |  2008  |  The Journal of Politics  |  |  |  |  |  |
| Arceneaux  |  2007  |  Quarterly Journal of Political Science  |  |  |  |  |  |
| Nickerson  |  2007  |  American Journal of Political Science  |  |  |  |  |  |
| Guan & Green  |  2006  |  Comparative Political Science  |  |  |  |  |  |
| Karp & Brockington  |  2005  |  The Journal of Politics  |  |  |  |  |  |
| Lassen  |  2005  |  American Journal of Political Science  |  |  |  |  |  |
| Lucinda  |  2005  |  Comparative Political Science  |  |  |  |  |  |
| Cardy  |  2005  |  The Annals of the American Academy of Political and Social Science  |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  Author(s)  |  Year  |  Source  |  Experimental?  |  Measure of Vote Choice  |  Measure of Turnout  |
|  |  |  |  |  Survey  |  Administrative  | Survey  |  Administrative  |
| Nickerson  |  2005  |  The Annals of the American Academy of Political and Social Science  |  |  |  |  |  |
| Ramirez  |  2005  |  The Annals of the American Academy of Political and Social Science  |  |  |  |  |  |
| Gerber  |  2004  |  American Behavioral Scientist  |  |  |  |  |  |
| Bélanger  |  2004  |  Comparative Political Science  |  |  |  |  |  |
| Green et al  |  2003  |  The Journal of Politics  |  |  |  |  |  |
| King  |  2001  |  The Journal of Politics  |  |  |  |  |  |
| Morgenstern & Zechmeister  |  2001  |  The Journal of Politics  |  |  |  |  |  |
| Gerber & Green  |  2000  |  American Political Science Review  |  |  |  |  |  |
| Gerber & Green  |  2000  |  The Journal of Politics  |  |  |  |  |  |
| Adams & Smith  |  1980  |  The Public Opinion Quarterly  |  |  |  |  |  |

**B** **Description of Treatment and Experiment**

In collaboration with a local NGO, the experiment involves the dissemination of information about incumbent legislative performance in advance of Benin’s April 26, 2015 National Assembly elections.[[1]](#footnote-1) The information provided was drawn from official reports of the Office of the President of the National Assembly.[[2]](#footnote-2) The video provided performance information about an incumbent legislator’s: 1) rate of attendance at legislative sessions, 2) rate of posing questions during legislative sessions, 3) rate of attendance in committees, and 4) productivity of committee work (the number of laws considered by the committee). The video provided raw data for each of these four performance indicators and presented two summary indicators. The first, an index of plenary performance on a scale of 1-10, took the average of normalized scores on the first two indicators: attendance and participation during full legislative plenary sessions. The second, an index of committee performance also on a scale of 1-10, took an average of the normalized scores on the second two indicators: attendance at committee meetings and productivity. To further synthesize the performance information, a global performance index which averaged scores from the first two indices was provided. In Appendix F, we discuss our efforts to validate these indices using a separate dataset on politician wealth and interviews with legislators.

In addition to randomizing whether voters received legislative performance information or not, our study design varied the content of the treatment message to vary the signal of the importance of the legislative performance dimension. Treated participants were shown a video with either *only* the information about relative legislator performance (*Info Only*), or that same information *plus* an additional message highlighting the importance of legislative performance to voter welfare (*Salience*). We also varied the method by which the information was disseminated. Treated participants received the intervention either privately by watching a video on a smartphone in the respondent’s household (*Private*) or publicly through the screening of the same video via a projector in a public location in the village or quarter (*Public*). In addition, we varied dosage — the density of treated villages in the commune. Participants were told during the intervention how many other villages in their commune were receiving legislative performance information. In high dosage communes, we randomly assigned 3 villages to each of the four combinations of content plus dissemination method (Info-Only/Private, Info-Only/Public, Salience/Private, Salience/Public), for a total of 12 villages treated with legislative performance information. In low-dosage communes, we assigned only one village to treatment (Civics/Public). In this paper, we report on the baseline treatment condition, Information Only in high-dosage communes.

Figure B.1 outlines the sampling and randomization procedure and shows each of the experimental conditions with the sample size of villages and survey respondents in each condition. While baseline and endline surveys were conducted in only 45 of the control villages, administrative data were collected on all polling stations in the sample communes to serve as a larger control group (N=1,071) for the analysis of electoral outcomes.

Figure B.1: CONSORT diagram

**C** **Randomization**

We implemented a two-stage randomization procedure. First, we randomly assigned each of the 30 communes in our sample to either the low or the high dosage condition (we only report effects in the high-dosage communes in this paper), blocking on incumbent legislative performance, which is observed at the commune level, and on north/south, since being in the culturally distinct north or south of the country is an important moderator of political behavior in Benin.[[3]](#footnote-3) Within 4 blocks (high and low performance in the north and south) of communes, we assigned half to high-dosage and half to low-dosage treatment.

Second, we randomly assigned treatment conditions within communes. The unit of randomization was the rural village or its equivalent urban quarter, the lowest level of social and territorial organization. In high dosage communes, we randomly assigned each village/quarter to one of five conditions: 1) Information Only/Private, 2) Information Only/Public, 3) Information + Salience/Private, 4) Information + Salience/Public, or 5) Control. We only report on the baseline treatment condition (Information Only) in this paper.

Three villages/quarters in each of the 15 high dosage communes were randomly assigned to one of the four treatment conditions, and the remainder villages/quarters in the commune were assigned to the control group. Thus, in the high dosage communes, we have a  factorial design with a pure control group. To increase statistical efficiency, we assigned villages to experimental conditions while stratifying on urban/rural status and electoral competitiveness of the village in the previous legislative election. Because in each high-dosage commune 12 villages were assigned to treatment and three to control, the sample reported in this paper comprises 225 villages.

**4** **Institutional Context and Focus on Single Incumbents**

 Benin’s is a unicameral legislative system, and National Assembly elections are held every five years. The country is divided into 12 departments with two legislative constituencies in each, for a total of 24 constituencies. Each of these 24 constituencies then comprises, on average, three communes. Arrondissements nest within these communes, and in turn villages (or their urban equivalent, quartiers) nest within these arrondissements and represent the lowest-level of administration in the country.The country’s 83 legislators are elected directly through a party-list proportional representation system: voters vote for party lists rather than individual legislators. Additionally, Benin’s constituencies are multi-member.

Yet in spite of this institutional context, voters pay greater attention to individual candidates than to political parties. This occurs for several reasons. First, even though constituencies are multi-member districts, in practice, many legislators focus on and “take care of” a particular commune within their constituency, facilitating a one-to-one correspondence of incumbent legislator to commune. Indeed, each constituency comprises an average 3.2 communes; in turn, voters elect an average 3.5 deputies per constituency. Second, experts evaluate the party system in Benin as fragmented and weak (Banégas 2003; Gazibo 2012). New parties and new coalitions form at almost every electoral cycle, and Benin’s last two presidents have both been independent politicians, unaffiliated with existing political parties.[[4]](#footnote-4) Pre-experiment focus groups in our study confirmed that villagers can name and agree on a single legislator as their incumbent representative. In sum, although the formal institutional context favors party rather than candidate voting, voters in Benin tend to vote for candidates.

We restricted our experimental sample to 30 communes for which we could verify a one-to-one correspondence of incumbent to commune, and in which the incumbent was running again. Figure D.1 shows the sample communes.

Figure D1: Sample Communes and Villages/Quarters

*Note:* The map displays all sample communes (in yellow) and sample villages and urban quarters.

**E** **Survey Sampling Procedures**

The sampling procedure for the baseline survey occurred as follows: enumerators used a random walk procedure to select compounds in which to administer the baseline survey and/or intervention. In each village, two enumerators started at the place where they had met with the chief to ask permission to survey (or at a central landmark in an urban quarter). They then spun a pen in a circle to select a direction in which to walk (or section of the quarter to cover, in the case of urban quarters), and followed the ends of the pen to walk in different directions and approach compounds at a set interval (determined by the number of compounds in the village/quarter). Within compounds, individual respondents were randomly selected from the list of adult members of the compound while alternating on gender. As a condition of participating in the baseline survey, respondents had to have access to a cell phone.[[5]](#footnote-5) Respondents were then recontacted by phone during the endline survey. A total of 3,419 individuals participated in the baseline *and* endline surveys (6,132 in the baseline), with an additional 6,174 receiving the intervention (or an invitation to the public screening) but no survey. To maximally harmonize public and private treatments, we endeavored to treat the same numbers of individuals per village across conditions. Thus, we provided the private treatment to 40 individuals in each private village even though we surveyed only a random half of those.

In each Private village, 20 people were randomly selected both to take the survey and to receive the intervention, 20 people were randomly selected only to receive the intervention, and 10 people were randomly selected to serve as control individuals and thus took the survey but received no intervention. In Public villages, 20 people were randomly selected to take the survey and be invited to receive the intervention at a public workshop, while an additional 40 people were randomly selected only to be invited to receive the intervention. We sampled in this way so as to ensure that roughly the same number of people would be treated in both Private and Public villages.[[6]](#footnote-6)

**F** **Validating the Performance Index**

To validate our performance measure, we examine whether our index correlates well with an alternative — and independently created — proxy for legislative performance: the legislator’s professional background prior to holding office. Exploring the rising cost of campaigns and the role of money in politics in Benin, Koter (2016) shows that wealthy individuals (business people and customs officials) have more than quadrupled their presence in parliament while the presence of the less wealthy, intellectual class (teachers, lawyers, academics) who comprised the vast majority of parliamentary seats early in Benin’s democracy has been steadily declining. While the latter politicians are considered better qualified to fulfill the formal duties of their position, the former are more valuable to parties because of their ability to buy votes. Combining our performance index with occupational data collected by Koter (2016), we see that wealthier politicians perform about 50 percent less well than other parliamentarians on components of the index such as attendance at plenary sessions and committee meetings. This increases our confidence that the performance index is measuring true legislative capacity.

More anecdotally, our elite interviews during an extraordinary session of parliament also revealed types consistent with our index. An example of a “good" performer we interviewed was a retired agronomist, who complained that he entered politics to address the concerns of his impoverished rural neighbors through legislation but was disappointed to learn that most politicians enter parliament to advance personal aims rather than the interests of the nation. Meanwhile, “bad" performers were difficult to interview because they were not even in the capital during the extraordinary session of parliament.[[7]](#footnote-7) In short, “good" performers according to our index were indeed politicians interested in lawmaking and who were active during an extraordinary session of parliament. By contrast, “bad" performers according to our index were notably absent from the capital, and in some cases, from the country.

**G** **Aggregating Official Data and Balance Tests**

We were able to match 2015 polling station data to all villages in our experimental sample except for one treated village and two surveyed control villages, which we drop from the analysis.[[8]](#footnote-8) Including control villages that were not surveyed, among all villages and quarters in our original sample of 30 communes, we were able to match 88% to the 2015 outcome data.[[9]](#footnote-9) On most pre-treatment characteristics unmatched and matched villages are statistically indistinguishable. See Table G.1.

Table G.1: Balance Across Villages Matched and Unmatched to Administrative Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean Unmatched | Mean Matched |  Difference |  P-Value |
| Registered Voters (log) |  6.34 |  6.53 |  .19 |  .09  |
| Urban  |  .28 |  .23 |  .06 |  .14  |
| Turnout  |  68.39 |  67.54 |  .86 |  .66  |
| Competitive (dichotomous) |  .52 |  .45 |  .07 |  .18  |
| Incumbent Performance |  4.98 |  5.19 |  .22 |  .69  |
| North  |  .39 |  .46 |  .07 |  .65  |
|  P-values generated from tests in which we cluster on commune. |

Table G.2: Balance Between High and Low Dosage Communes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean Coordination e | Mean No Coordindation |  Difference |  P-Value |
| Registered Voters (log) |  6.3 |  6.52 |  .22 |  .18 |
| Urban  |  .21 |  .25 |  .04 |  .41 |
| Competitive (dichotomous) |  .42 |  .5 |  .07 |  .25 |
| Vote Margin  |  .28 |  .24 |  .05 |  .19 |
| Overall Performance |  4.97 |  5.35 |  .38 |  .67 |
|  P-values generated from tests in which we cluster on commune. |

Table G.3: Balance in High Dosage Communes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  Control | Info Only/Private | Info Only/Public | Salience/Private | Salience/Public |
| Registered Voters (log) |  687.5 |  828.26 |  1066.25 |  1110.73 |  807.27 |
|  |  |  (.08)  |  (.06)  |  (.00) |  (.02) |
| Urban  |  .18 |  .03 |  .34 |  .27 |  .29 |
|  |  |  (.00) |  (.00) |  (.07) |  (.01) |
| Competitive (dichotomous) |  .41 |  .49 |  .48 |  .5 |  .48 |
|  |  |  (.11) |  (.30) |  (.14) |  (.25) |
| Vote Margin  |  .29 |  .25 |  .23 |  .26 |  .31 |
|  |  |  ( .14) |  (.05) |  (.44) |  (.47) |
| Overall Performance |  4.89 |  5.26 |  5.19 |  5.23 |  5.23 |
|  |  |  (.14) |  (.21) |  (.17) |  (.17) |
|  P-values in parantheses indicate significance of difference between the mean and each treatment group and the control group mean.  |
|  P-values generated from tests in which we cluster on commune. |

**Note:** Because of our blocking and randomization process, there is a lack of balance in the raw means on urban and number of registered voters. This occurred because our rural blocks, where there are also fewer registered voters, contain larger numbers of units than our urban blocks. Since all non-treated units are used as controls, on average the proportion of rural areas in control is lower than in treatment. This lack of balance is not a problem as we use block fixed effects in all of our analyses, which controls for the urban/rural difference.

Table G.4: Balance in Low Dosage Communes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean Treatment | Mean Control |  Difference |  P-Value |
| Registered Voters (log) |  933.27 |  1024.24 |  90.97 |  .76 |
| Urban  |  .27 |  .25 |  .01 |  .92 |
| Competitive (dichotomous) |  .53 |  .5 |  .04 |  .78 |
| Vote Margin  |  .25 |  .24 |  .01 |  .86 |
| Overall Performance |  5.42 |  5.35 |  .07 |  .9 |

**H** **Defining Good and Bad News**

We expect voters to respond differently to information about legislators that is positive and to information that is negative (though we note here and specify below that our analysis is intra-legislator: we compare villages that do and do not receive information about the same legislator). Since the information provided explicitly compares the incumbent legislator’s performance to the performance of legislators in the surrounding area (those in the same department), we code positive and negative information relative to this local benchmark. More specifically, we define the information as positive if the incumbent’s overall score is better than that of other deputies in the department. Poor legislative performers are those whose overall legislative score is worse than that of other legislators in this local area. This coding rule was pre-specified in our pre-analysis plan.[[10]](#footnote-10)

**I** **Model Specification**

As pre-specified, we divide the sample into communes where the incumbent is a strong legislative performer and communes where the incumbent is a poor legislative performer and run analyses separately. To analyze treatment effects of receiving good [bad] news about one’s incumbent, we estimate the following model using OLS

  (1)

where  represents the vote share of the incumbent deputy’s party in village  of commune , and  represents a complete set of block fixed effects. Treatment is assigned within three sub-blocks within each commune (rural-competitive, rural-noncompetitive, and urban). There are therefore 45 blocks (15 communes x 3 sub-commune blocks). Our use of block fixed effects ensures that our experimental estimates are being driven by comparisons of similar villages/quarters within the same commune and *with the same incumbent.*

 is a generic treatment variable indicating the treatment status of village  in commune . We cluster standard errors by commune-treatment condition. We pre-specified that we would run all analyses with and without additional pre-treatment covariates. Because the results are substantively unchanged with covariates and because we lose the majority of our sample when we include covariates from the survey, we present results without controls.

1. In designing the experiment, we followed the ethical principles agreed upon by the Metaketa initiative, as outlined in the joint metaketa pre-analysis plan: that the intervention consist of information that existed in the political system, be provided with consent, in a non-partisan way, without deception, and in cooperation with a local group. [↑](#footnote-ref-1)
2. Reports are supposed to be made publicly available but, in practice, are difficult to obtain. [↑](#footnote-ref-2)
3. See, e.g., Adida (2015). [↑](#footnote-ref-3)
4. Benin’s former president, Boni Yayi, first ran as an independent. Once he became President, a new coalition – the FCBE – formed in his support. It has since become a political party, and currently holds the most seats in the National Assembly. [↑](#footnote-ref-4)
5. According to the most recent round of Afrobarometer surveys in Benin (2014), approximately 77% of Beninois own a cell phone (http://www.afrobarometer.org). Still, respondents were not required to own a cell phone to participate in our intervention. The cell phone to which they had access could belong to a friend or relative.
 [↑](#footnote-ref-5)
6. On average, 55 individuals in Public villages attended the video screenings (range from 20 to 70), indicating a reasonable balance of treated individuals across Public and Private villages. [↑](#footnote-ref-6)
7. Of the parliamentarians interviewed, only one was a bad performer, and we had to travel to his home constituency as he does not typically attend parliamentary sessions. Other reasons we were unable to interview bad politicians included the legislator’s simple refusal to participate, business travel to Niamey or Brussels, our inability to locate the legislator, or the legislator’s lack of fluency in French. [↑](#footnote-ref-7)
8. In one village, enumerators administered both private screenings and a public screening, although in both cases they showed the same video (Info-Only). We thus exclude this village from our analyses comparing public and private conditions. [↑](#footnote-ref-8)
9. A number of quarters (the urban administrative equivalent of a village) were split into two between 2011 and 2015. We were able to match some of these with the help of our local partner, but not all. [↑](#footnote-ref-9)
10. Since we do not have information on voter priors in non-surveyed control villages, we leverage the fact that voters in Benin are generally quite uninformed about the legislative performance of their incumbents. In our baseline survey, 54% of participants report they “do not know” whether their incumbent’s legislative performance is better than that of other deputies in the local area. Additionally, interviews with Beninese radio hosts confirm that voters have no access to this type of information, and therefore no way of forming such beliefs. We therefore treat voters as having highly diffuse priors about incumbent legislative performance, allowing us to define positive and negative information based solely upon the information provided in the experiment. [↑](#footnote-ref-10)