Online Appendix for "Pretending to Support? Duterte's Popularity and Democratic Backsliding in the Philippines"

(Figure A.6 has been corrected and replaced.)

Contents

A. Validity of Assumptions **2**

A.1. Test for the no design effect assumption 2

A.2. Discussion of the no liars assumption 2

A.3. Additional combined tests of several assumptions 4

B. Robustness Checks **5**

B.1. Direct question-based estimates based only on the control group 5

B.2. Treating "don't know" responses as "disapprove" 6

B.3. Estimates by combining list experimentation and direct questioning 8

B.4. Supplementary information for our sensitivity analysis for non-strategic misreporting 10

C. Supplementary Information **14**

C.1. Details of the exploratory subgroup analyses 14

C.2. Cross-tabulation of various subgroup combination 18

C.3. Subgroup analyses by ethnicity 20

References **21**

A. Validity of Assumptions

A.1. Test for the no design effect assumption

List experiments require us to assume that there are no design effects, which means that adding a sensitive item to a list does not affect the sum of affirmative answers to the remaining items. One of the scenarios in which this assumption does not hold is when "respondents evaluate control items relative to the sensitive item" (Blair and Imai 2012, 63). Some may suspect that this scenario applies to this study and that respondents assigned to the treatment group considered whether they approve of Quezon, Marcos, and Aquino by comparing them to Duterte.

Following Blair and Imai (2012), we examined all proportions of respondent types estimated to be positive. Respondent types, notated by π_{yz}, refer to the pairs of each respondent's potential affirmative count for control items (denoted by y) and potential response to the sensitive item (denoted by z). We also conducted Blair and Imai's (2012) diagnostic test for the no design effects assumption, which is implemented by `ict.test()` in R package `list` (Blair and Imai 2010).

Table A.1 shows the results. All respondent types produced a positive proportion estimate. The diagnostic test did not reject the null hypothesis that $\pi_{yz} \geq 0$ for all y and z (p-values were 1.000 for both face-to-face and online surveys).

A.2. Discussion of the no liars assumption

We expect that respondents assigned to the treatment group answered the list questions based on their true applicability to the sensitive item (i.e., true preference for Duterte). This assumption is called the no liars assumption (Blair and Imai 2012). Although the difference-in-means estimator employed in this study requires a weaker assumption (Ahlquist 2018), previous studies conducting list experiments often made the no liars assumption. As such, we follow them.

Given that choosing "RODRIGO 'DIGONG' DUTERTE" in the list is regarded as "socially desirable" in this study, an important sign of this assumption's violation is a non-negligible proportion of respondents in the control group answering "0" for the list question (i.e., many respondents did not support any politician in the control list). This is because if respondents

Table A.1. Estimated proportions of respondent types

	Face-to-Face		Online	
	Est.	S.E.	Est.	S.E.
$\hat{\pi}_{01}$	0.048	0.010	0.019	0.007
$\hat{\pi}_{11}$	0.230	0.028	0.137	0.014
$\hat{\pi}_{21}$	0.132	0.022	0.200	0.015
$\hat{\pi}_{31}$	0.097	0.012	0.170	0.008
$\hat{\pi}_{00}$	0.007	0.003	0.042	0.004
$\hat{\pi}_{10}$	0.333	0.022	0.176	0.011
$\hat{\pi}_{20}$	0.140	0.026	0.191	0.016
$\hat{\pi}_{30}$	0.013	0.018	0.065	0.013

Note: $\hat{\pi}_{yz}$ indicates the estimated proportions of respondent types, where y represents potential affirmative counts for the control items, and z represents a dummy variable for the potential response to the sensitive item (one represents affirmative for the sensitive item).

assigned to the treatment group did not respond affirmatively to all of the list items, their sincere responses (i.e., "0") would reveal that they did not support Duterte. Due to fear of exposing their "socially undesirable" preference for Duterte, such respondents would lie when answering the list question; this is called the problem of floor effects (Blair and Imai 2012).

Figure A.1 shows the distribution of the outcome variable of the list question. The proportion of control-group respondents who answered "0" was small for both survey types. Specifically, the relevant percentage did not exceed 6% for either survey, a value comparable to or even smaller than the corresponding value in list experiments conducted by previous studies (see Supplementary Information C of Kuhn and Vivyan 2022 and the studies cited there). As such, we can consider that a floor effect was not a serious problem in this study.

However, Figure A.1 reveals that many respondents in the control group chose the maximum value "3" (specifically, 11% in the face-to-face survey and 24% in the online survey). This may provoke suspicion about a ceiling effect, which means that "the respondents in the treatment group whose truthful answer is [negative] only for the sensitive item may fear that their true preference for the sensitive item would be revealed by giving ['4'] as their answer" (Blair and Imai 2012, 66). However, because, unlike in the case of a floor effect, answering "4" would not directly disclose respondents' disapproval of Duterte, the problem of ceiling effects is less

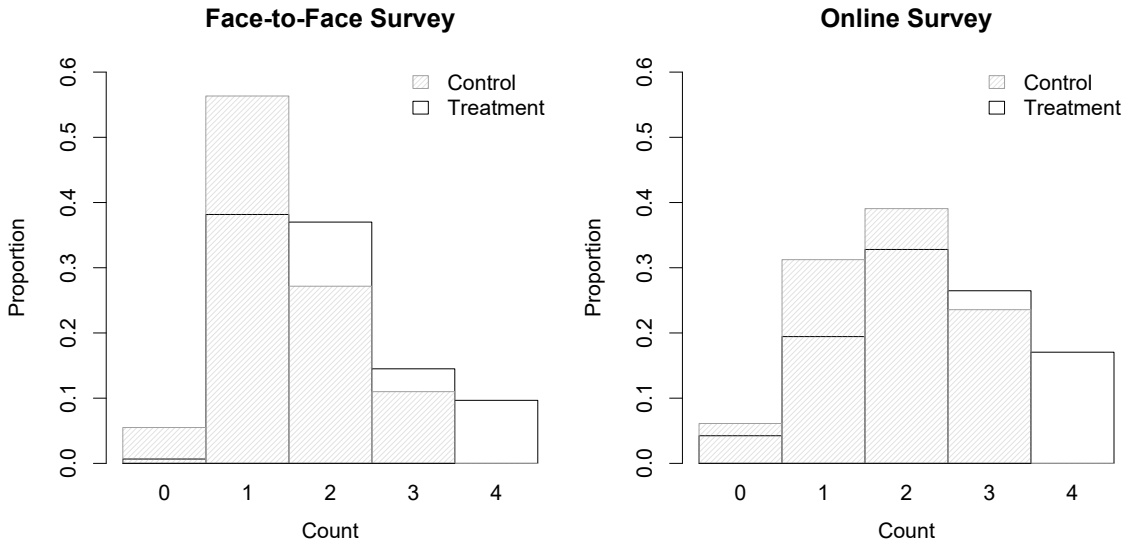


Figure A.1. Distribution of the outcome variable of the list question

critical for our experiment.

A.3. Additional combined tests of several assumptions

We also conducted two placebo tests proposed by Aronow et al. (2015). The first test simultaneously checks the validity of several assumptions: monotonicity, no liars, no design effects, and treatment independence. The test utilizes the fact that under these assumptions, the difference in responses to the list question between the treatment and control groups is expected to be zero among those who sincerely revealed their disapproval of Duterte in the direct question. We used `combinedListDirect()` in `list` package to conduct this test. Because this function supposes that an affirmative answer to the sensitive item is socially undesirable, we reversed the coding of the response variables of the list and direct questions before running the function.

We found that in the face-to-face survey, the list-based estimate of Duterte's approval rate among those who revealed their disapproval in the direct question was 0.138, which was statistically indistinguishable from zero ($p = 0.375$). In contrast, the corresponding estimate in the online survey was 0.211, which is significant at the 1% level ($p = 0.007$).

These results mean that there is no evidence any of the four assumptions listed above were violated in the face-to-face survey but that one or more assumptions could have been violated in

the online survey. As discussed in Section 4.4 in the main text, our original placebo experiment suggested that there were non-strategic misreporters in the online survey, which should have violated the assumptions. Therefore, our estimate of the absolute level of Duterte's approval rate found in the online survey may be unreliable. Despite this fault, we argue that our conclusion that the SDB existed holds under the wide range of parameters regarding the percentage of non-strategic misreporters.

The second Aronow et al.'s (2015) placebo test checks the treatment independence assumption. Given that random assignment is successful, this assumption is violated if "the treatment assignment affects the response to the direct question" (Aronow et al. 2015, 47), which was the scenario that concerned Blair and Imai (2012). The placebo test simply examines whether the difference in the means of responses to the direct question exceeds zero.

We conducted this test using the same function mentioned above and found that the null hypothesis was not rejected (p-values were 0.952 and 0.828 for the face-to-face and online surveys), which implies that there is no evidence of a violation of the treatment independence assumption.

B. Robustness Checks

B.1. Direct question-based estimates based only on the control group

Blair and Imai (2012) recommended not using data from the treatment group to obtain direct question-based estimates of the concerning prevalence because a list experiment question may prime respondents in the treatment group, distorting their responses to the direct question. Assuming the effect of such priming would be small, we decided to utilize data from the treatment group to improve estimation efficiency in accordance with our preregistration. The result of the second Aronow et al.'s (2015) placebo test on the treatment independence assumption reported in Section A.3 supports our decision to use the entire sample to efficiently estimate the direct question-based approval rate.

Nonetheless, we retested our hypotheses with the direct question-based estimates of Duterte's approval rating based only on the control group. Figures A.2 and A.3 show the results. We can

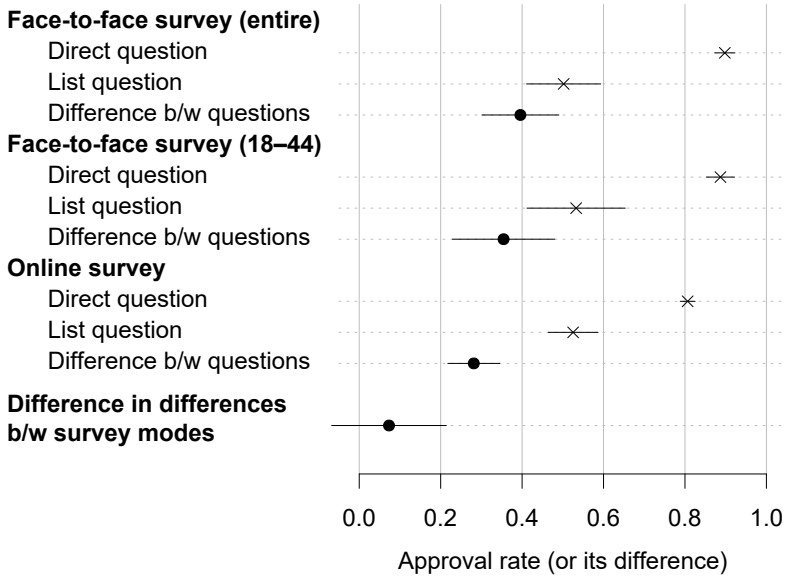


Figure A.2. Estimates of Duterte's approval ratings, SDB, and difference in SDB based on DQ-based estimates using only the control group

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

confirm that these figures are almost identical to Figures 2 and 3 in the main text, indicating that our conclusions hold.

B.2. Treating "don't know" responses as "disapprove"

In the main analysis, we treated "Don't know" (DK) responses to the direct question (DQ) as missing values, assuming that the DQ responses were randomly missing. However, we cannot empirically verify that this assumption holds. In this section, we assume that all respondents who chose DK did not approve of Duterte. That is, we define the DQ-based estimate of approval rating as the percentage of "Approve" responses, where the denominator is the total number of respondents (not excluding DK respondents). This definition yields the most conservative estimate of the SDB.

Figures A.4 and A.5 show the results of the replication of our main analysis using this alternative definition of the DQ-based estimate of approval rating. Although the DQ-based estimates are lower than the original estimates, this does not change our conclusion that Hypotheses 1 and 6 hold and that the remaining hypotheses were not supported.

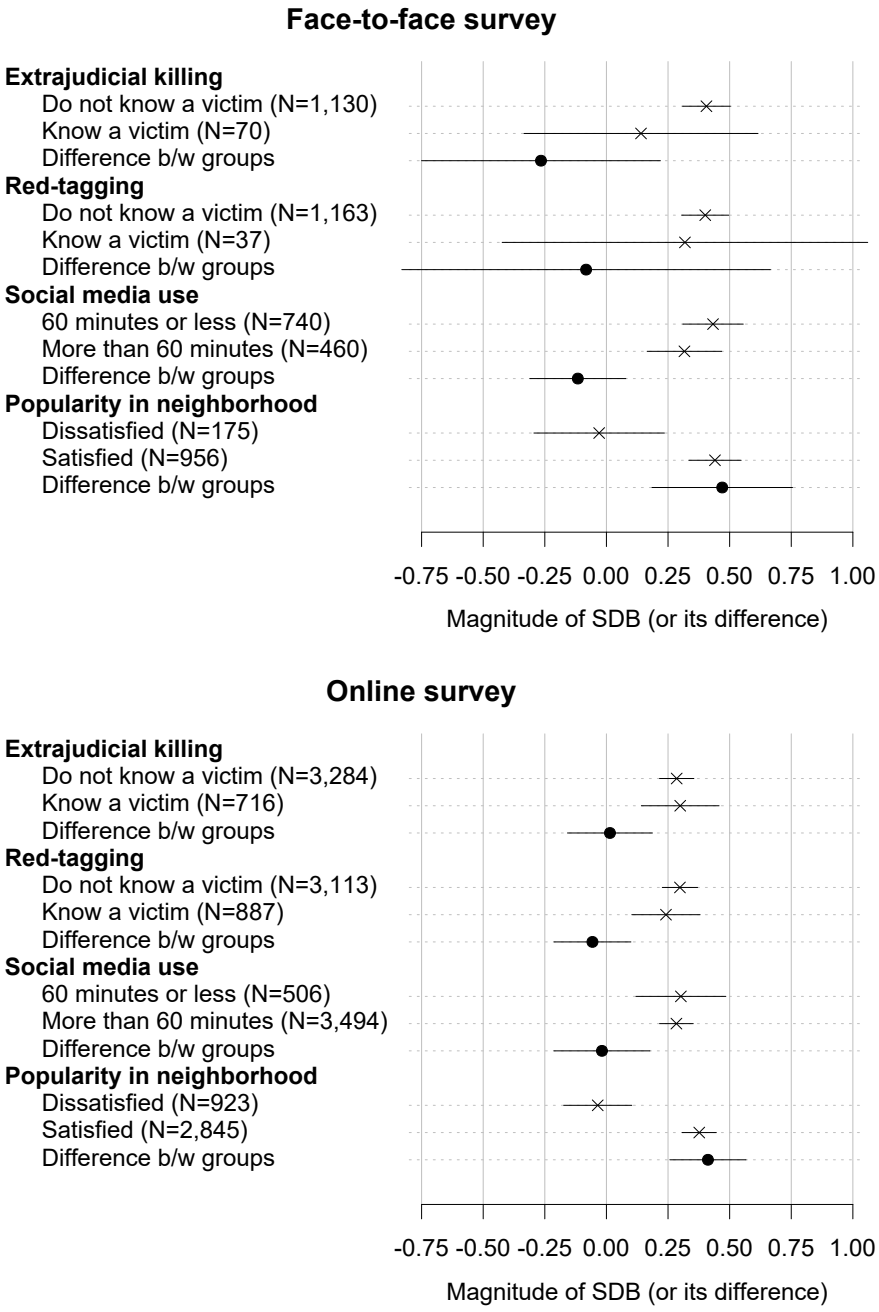


Figure A.3. Subgroup comparison of the magnitudes of SDB based on DQ-based estimates using only the control group

Note: Dots represent point estimates, and horizontal segments represent the 95% CIs.

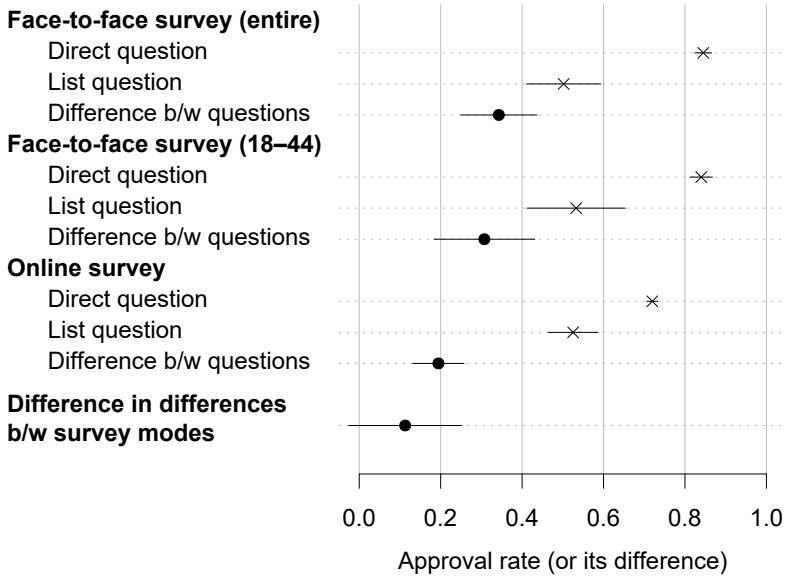


Figure A.4. Estimates of Duterte's approval ratings, SDB, and difference in SDB based on an alternative definition of DQ-based estimates

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

B.3. Estimates by combining list experimentation and direct questioning

Because we directly asked all respondents, including those who were assigned to the treatment group, about their support for Duterte, we can employ Aronow et al.'s (2015) estimator, which uses responses to both the list experiment and the direct question. Aronow et al. (2015) argued that this estimator is more efficient than the standard difference-in-means estimator. As a robustness check, our data was reanalyzed relying on their estimator. We used `combinedListDirect()` in `list` package. For technical reasons, we did not consider block-specific intercepts as we did in the main analysis.

Figures A.6 and A.7 show the results, which are largely similar to the results reported in the main text. The only exception is that in Figure A.7, SDB among online respondents who perceived their neighbors to be dissatisfied with Duterte was estimated to be higher than the original estimate, which results in shrinkage of the difference in SDB between "dissatisfied" and "satisfied" groups. Nonetheless, the difference is still statistically significant, and Hypothesis 6 is supported.

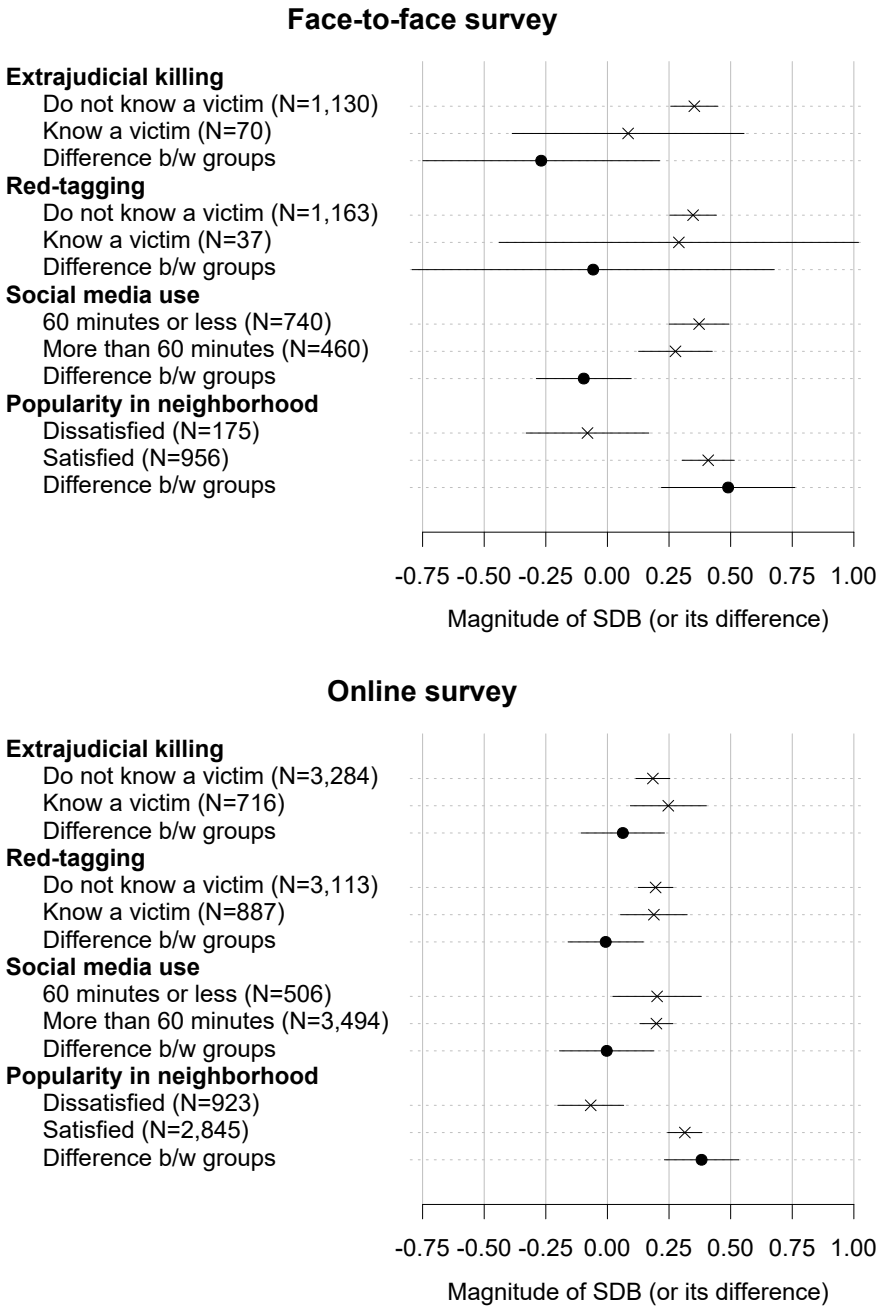


Figure A.5. Subgroup comparison of SDB magnitudes based on an alternative definition of DQ-based estimates

Note: Dots represent point estimates, and horizontal segments represent the 95% CIs.

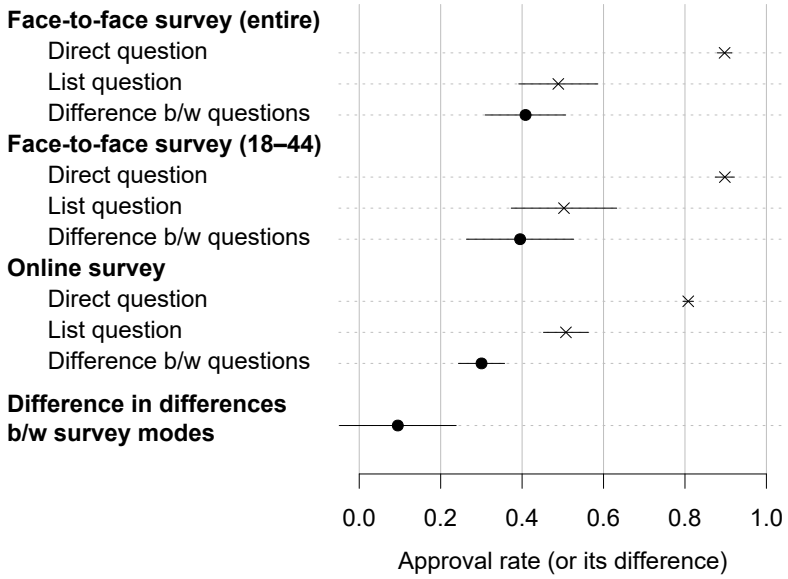


Figure A.6. Estimates of Duterte's approval ratings, SDB, and difference in SDB using Aronow et al.'s (2015) estimator

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

B.4. Supplementary information for our sensitivity analysis for non-strategic misreporting

B.4.1. Difference from existing design

Our placebo survey design to conduct sensitivity analysis for non-strategic misreporting is different from Riambau and Ostwald's (2021) survey design, though they also use the term "placebo." Riambau and Ostwald (2021) proposed that scholars add a placebo statement not expected to apply to any individual to the control group's list to make it equal in length to the treatment group's list. However, in their supplemental information, Kuhn and Vivyan (2022) recognized that their design does not eliminate the non-strategic misreporting bias, although it can fix the sign of the bias (see also Agerberg and Tannenberg 2021). Differing from Riambau and Ostwald's (2021) design, we included a placebo item in a list for the treatment group and did not include an additional item in a list for the control group, meaning that list lengths were not equal across groups.

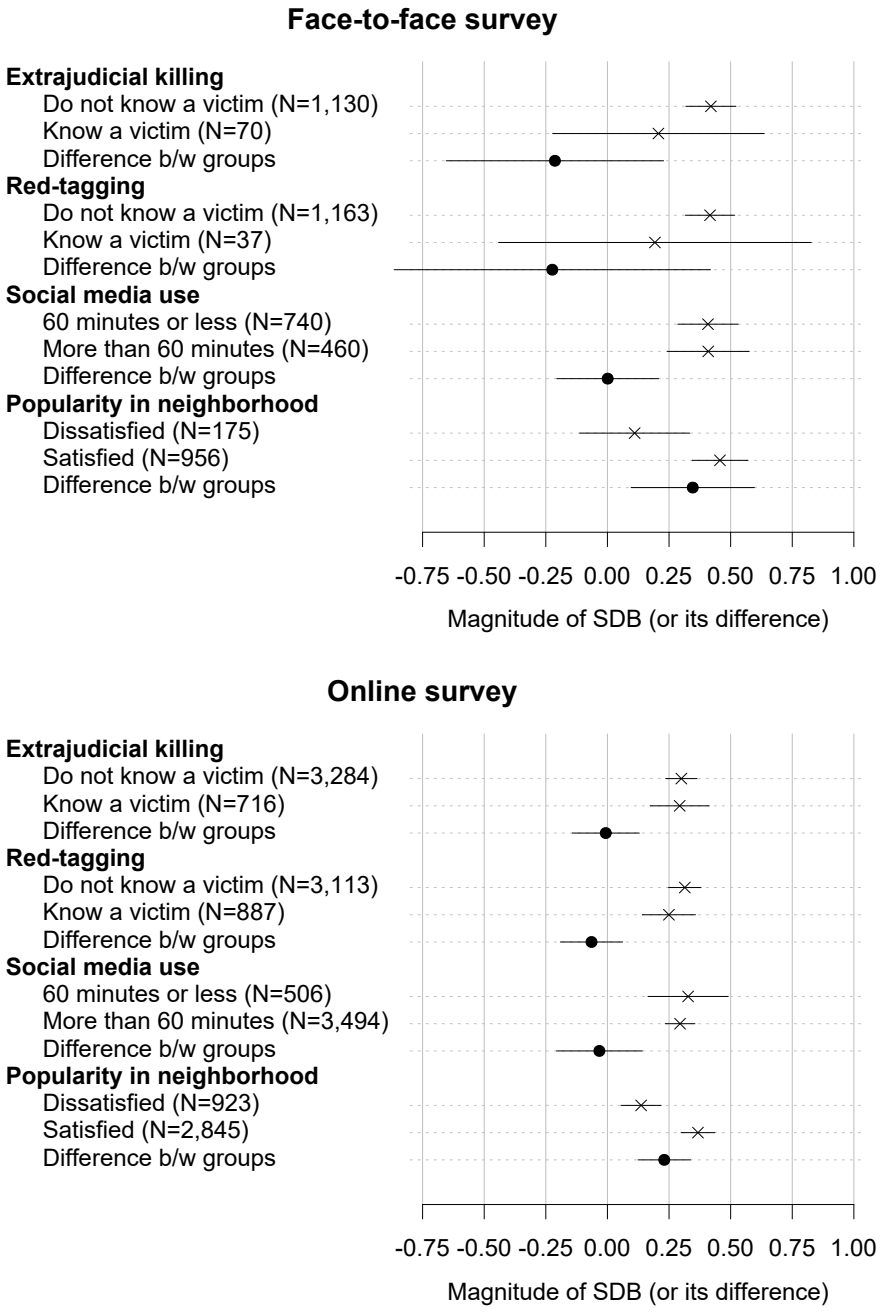


Figure A.7. Subgroup comparison of SDB magnitudes using Aronow et al.'s (2015) estimator

Note: Dots represent point estimates, and horizontal segments represent the 95% CIs.

B.4.2. How our placebo survey design works

Our strategy's function is based on Kuhn and Vivyan's (2022) formulation of non-strategic error in Supplemental Information I. Following their notation, π is the true prevalence of a sensitive item, $\hat{\pi}$ is the difference-in-means estimate of π in a list experiment, λ is the percentage of non-strategic misreporters, and S^* is a binary indicator for non-strategic misreporter status. The expectation of $\hat{\pi}$ is as follows:

$$\mathbb{E}(\hat{\pi}) = (1 - \lambda)\pi + \lambda\mathbb{E}(\hat{\pi}_{S^*=1}), \tag{A.1}$$

where $\mathbb{E}(\hat{\pi}_{S^*=1})$ is the expected difference in means between non-strategic misreporters. This equation indicates that $\hat{\pi}$ is biased unless there are no non-strategic misreporters ($\lambda = 0$) or $\mathbb{E}(\hat{\pi}_{S^*=1})$ happens to equal π.

Below we use superscript (M) and (P) to denote the main (online) and placebo list experiment, respectively. Because the main (online) and placebo surveys are completely equivalent, we can plausibly assume that $\lambda^{(\mathrm{M})} = \lambda^{(\mathrm{P})}$ and $\mathbb{E}(\hat{\pi}^{(\mathrm{M})}_{S^*=1}) = \mathbb{E}(\hat{\pi}^{(\mathrm{P})}_{S^*=1})$. Therefore, Equation (A.1) for the main list experiment can be reduced to the following:

$$\begin{aligned}
\mathbb{E}(\hat{\pi}^{(\mathrm{M})}) &= (1 - \lambda^{(\mathrm{M})})\pi^{(\mathrm{M})} + \lambda^{(\mathrm{M})}\mathbb{E}(\hat{\pi}^{(\mathrm{M})}_{S^*=1}), \\
&= (1 - \lambda^{(\mathrm{M})})\pi^{(\mathrm{M})} + \lambda^{(\mathrm{P})}\mathbb{E}(\hat{\pi}^{(\mathrm{P})}_{S^*=1}), \\
&= (1 - \lambda^{(\mathrm{M})})\pi^{(\mathrm{M})} + \mathbb{E}(\hat{\pi}^{(\mathrm{P})}) - (1 - \lambda^{(\mathrm{P})})\pi^{(\mathrm{P})}, \\
&= (1 - \lambda^{(\mathrm{M})})\pi^{(\mathrm{M})} + \mathbb{E}(\hat{\pi}^{(\mathrm{P})}) - (1 - \lambda^{(\mathrm{M})})\pi^{(\mathrm{P})}.
\end{aligned} \tag{A.2}$$

Consequently, assuming that $\lambda^{(\mathrm{M})}$ is not equal to one, the quantity of interest, $\pi^{(\mathrm{M})}$, is expressed as follows:

$$\pi^{(\mathrm{M})} = \frac{\mathbb{E}(\hat{\pi}^{(\mathrm{M})}) - \mathbb{E}(\hat{\pi}^{(\mathrm{P})})}{1 - \lambda^{(\mathrm{M})}} + \pi^{(\mathrm{P})}. \tag{A.3}$$

When we substitute the difference-in-means estimates obtained in both surveys for their expectations, we know the numerator of the first term of the right-hand side of Equation (A.3). Moreover, in our placebo list experiment, we can consider that $\pi^{(\mathrm{P})}$ is close to zero.

Although $\lambda^{(\mathrm{M})}$ is unobserved (and unobservable), we can simulate how our prevalence

estimate for the sensitive item "RODRIGO 'DIGONG' DUTERTE" in the main list experiment is susceptible to the proportion of non-strategic misreporters by computing Equation (A.3) with varying $\lambda^{(\mathrm{M})}$. Additionally, we essentially expect that no person approved of a fictitious politician's performance (i.e., $\pi^{(\mathrm{P})} = 0$), but we cannot deny the possibility that some (even non-satisficing) respondents mistook "Paul Nueva" for a local politician with a name similar to a politician in their community. Thus, we also varied $\pi^{(\mathrm{P})}$ in a range close to zero in the simulation to examine its influence on our estimation of $\pi^{(\mathrm{M})}$.

B.4.3. Plausible lower limit of the proportion of non-strategic misreporters

Here we assume that the approval rate of "Paul Nueva" is zero (i.e., $\pi^{(\mathrm{P})} = 0$). Letting p denote the percentage of non-strategic misreporters in the sample of our placebo survey, we rewrite Equation (A.1) before taking the expectation for the placebo list experiment and obtain:

$$\hat{\pi}^{(\mathrm{P})} = (1-p)\hat{\pi}^{(\mathrm{P})}_{S^*=0} + p\hat{\pi}^{(\mathrm{P})}_{S^*=1}, \tag{A.4}$$

$$\therefore\ p = \frac{\hat{\pi}^{(\mathrm{P})} - \hat{\pi}^{(\mathrm{P})}_{S^*=0}}{\hat{\pi}^{(\mathrm{P})}_{S^*=1} - \hat{\pi}^{(\mathrm{P})}_{S^*=0}}. \tag{A.5}$$

Now $\hat{\pi}^{(\mathrm{P})}_{S^*=1}$ and $\hat{\pi}^{(\mathrm{P})}_{S^*=0}$ indicate the difference in means between non-strategic misreporters and other respondents, respectively. To obtain Equation A.5, we reject the highly unlikely case that $\hat{\pi}^{(\mathrm{P})}_{S^*=0}$ incidentally equaled to $\hat{\pi}^{(\mathrm{P})}_{S^*=1}$.

$\hat{\pi}^{(\mathrm{P})}_{S^*=0}$ and $\hat{\pi}^{(\mathrm{P})}_{S^*=1}$ are unobservable because we cannot identify S^* for each respondent. Nonetheless, corresponding to the assumption that the approval rate of "Paul Nueva" is zero, we substitute zero to $\hat{\pi}^{(\mathrm{P})}_{S^*=0} = 0$.[1] Therefore,

$$p = \frac{\hat{\pi}^{(\mathrm{P})}}{\hat{\pi}^{(\mathrm{P})}_{S^*=1}}. \tag{A.6}$$

Moreover, we argue that we can reasonably assume that $0 < \hat{\pi}^{(\mathrm{P})}_{S^*=1} \leq 1$. $\hat{\pi}^{(\mathrm{P})}_{S^*=1}$ "depends on

[1] We might even consider the uncertainty of $\hat{\pi}^{(\mathrm{P})}_{S^*=0}$. However, because the scenario that $\hat{\pi}^{(\mathrm{P})}_{S^*=1} = 1$, which we discuss immediately below, is still quite unlikely and leads to a conservative assumption, we believe that it is too conservative to consider any further uncertainty.

the decision rule that nonstrategic misreporters use to pick their reported item count, and how the resulting reported item count varies as a function of whether they are asked about the longer treatment list or shorter control list" (Kuhn and Vivyan 2022, 384). One plausible decision rule is that respondents randomly choose a possible option with equal probability, which Ahlquist (2018) called errors caused by this mechanisms "uniform error." Albeit not equal probability, if non-strategic misreporters do some counting, the count of the short control list should not exceed that of the long treatment list (i.e., $0 \leq \hat{\pi}^{(\mathrm{P})}_{S^*=1}$), nor should the difference in count between the lists exceed the difference in the number of items (i.e., $\hat{\pi}^{(\mathrm{P})}_{S^*=1} \leq 1$).[2]

Our placebo survey revealed that $\hat{\pi}^{(\mathrm{P})}$ was 0.242. Under the above assumptions, we can deduce from Equation (A.6) that the possible lowest value for $\lambda^{(\mathrm{P})}$ is $\hat{\pi}^{(\mathrm{P})} = 0.242$.

B.4.4. Analysis broken down by educational level

We re-conducted the sensitivity analysis by dividing respondents by educational level. The threshold between low and high is the same as in the analysis in Section 4.3 in the main text.

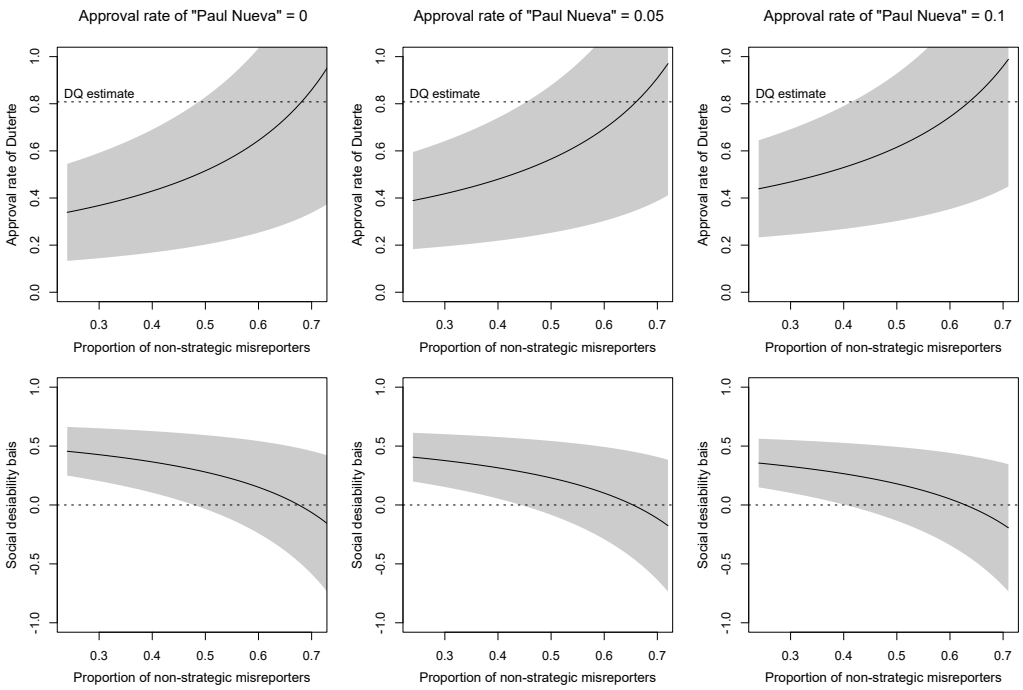
Figure A.8 shows the results. Splitting the sample did make the confidence intervals fairly wide. Nonetheless, when we focus on the point estimates, even for low-educated respondents, we can maintain the conclusion that the SDB existed unless we suppose that at least half the respondents were non-strategic misreporters.

C. Supplementary Information

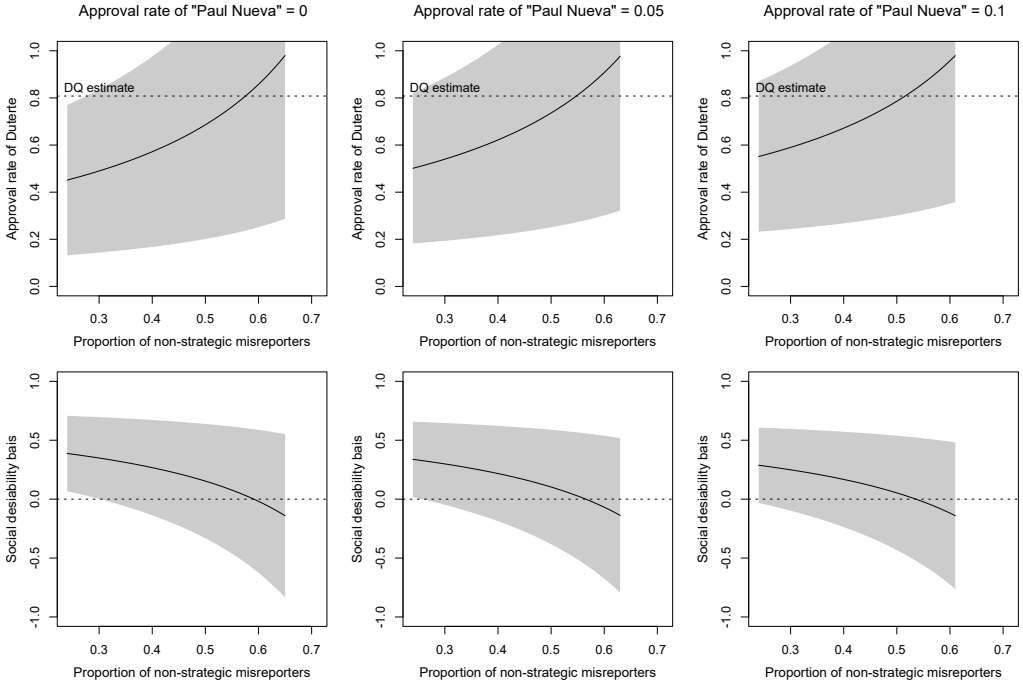
C.1. Details of the exploratory subgroup analyses

Figure A.9 shows DQ- and list-based estimates of Duterte's approval ratings and their differences (i.e., SDB estimates) compared across gender, education, and socioeconomic class. Figure A.10 is the same figure for comparison across region of residence.

[2] Because we rejected the case that $\hat{\pi}^{(\mathrm{P})}_{S^*=0} = \hat{\pi}^{(\mathrm{P})}_{S^*=1}$ and supposed that $\hat{\pi}^{(\mathrm{P})}_{S^*=0} = 0$, this argument leads the possible range of $\hat{\pi}^{(\mathrm{P})}_{S^*=1}$ to $0 < \hat{\pi}^{(\mathrm{P})}_{S^*=1} \leq 1$, not $0 \leq \hat{\pi}^{(\mathrm{P})}_{S^*=1} \leq 1$.



(a) Highly educated respondent



(b) Less educated respondent

Figure A.8. Results of the sensitivity analysis for non-strategic misreporting using a placebo list experiment broken down by educational level

Note: A solid line represents point estimates, and a shaded area represents the 95% CIs. A dotted line in the top row panels represents the direct question-based estimate of Duterte's approval rate in the main online survey.

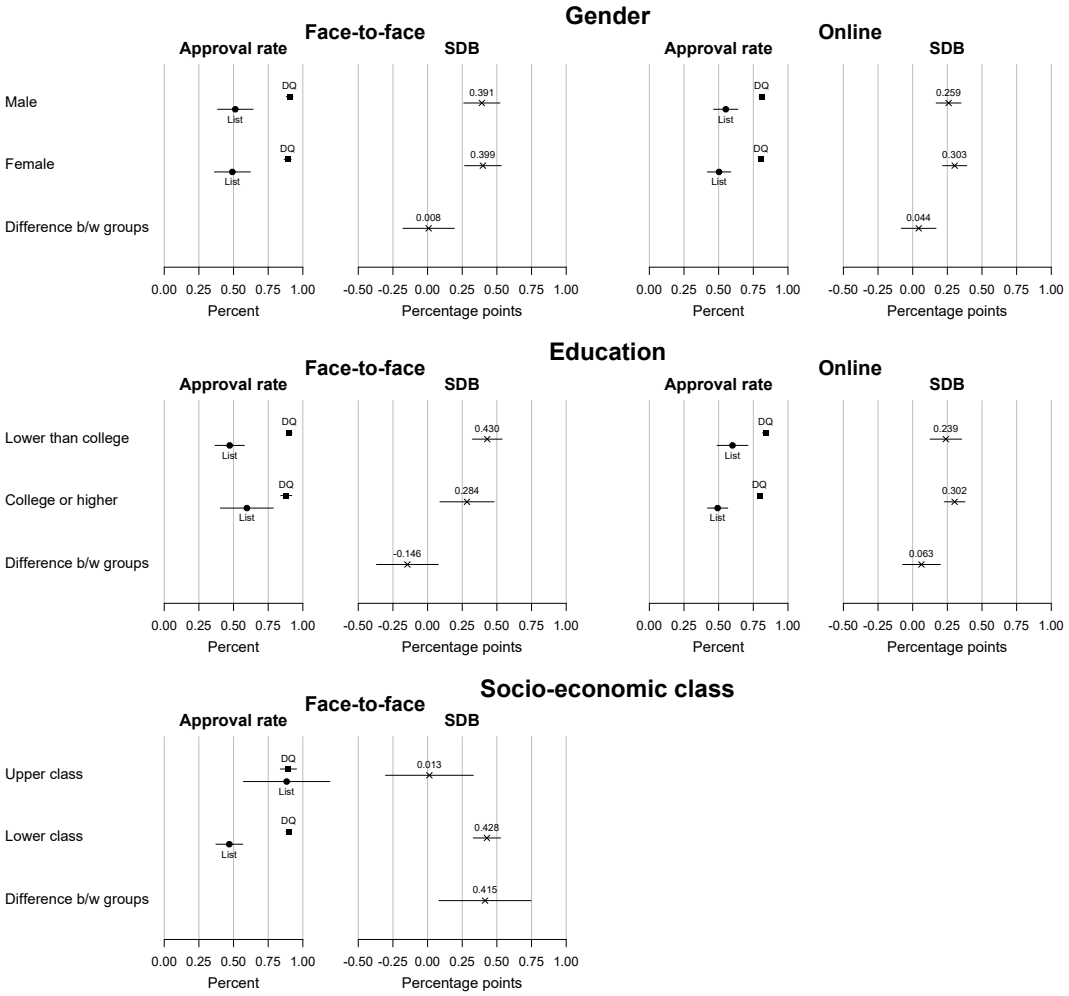


Figure A.9. Estimates of Duterte's approval ratings and SDB compared across demographic subgroups (gender, education, and socioeconomic class)

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

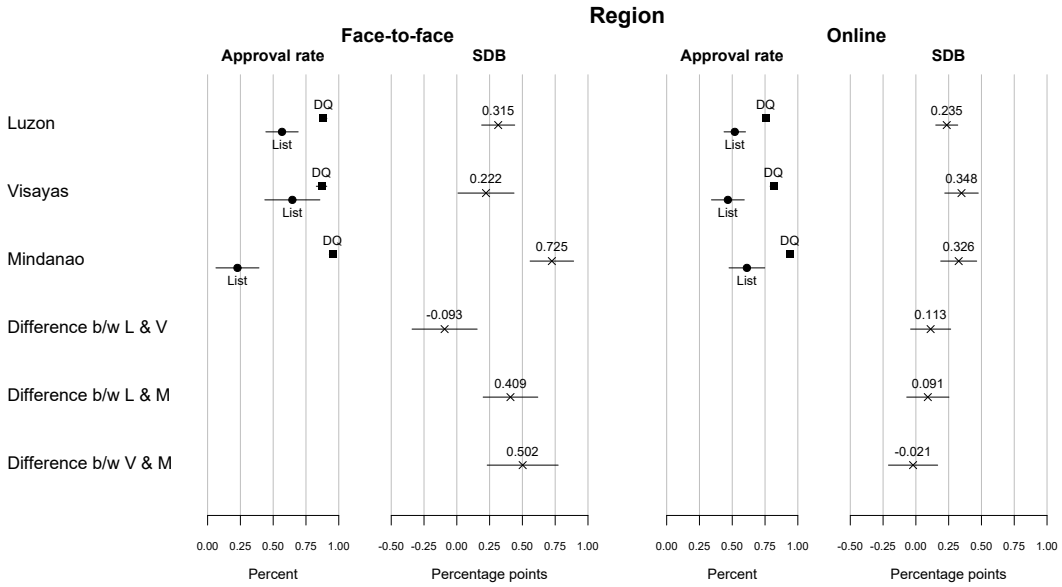


Figure A.10. Estimates of Duterte's approval ratings and SDB across demographic subgroups (region)

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

C.2. Cross-tabulation of various subgroup combination

Tables A.2 to A.7 report the cross-tabulation of various subgroup combinations.

Table A.2. Region and socioeconomic class

	ABC	DE	Total
Luzon	84	516	600
	(14.0%)	(86.0%)	(100%)
Visayas	15	285	300
	(5.0%)	(95.0%)	(100%)
Mindanao	14	286	300
	(4.7%)	(95.3%)	(100%)
Total	113	1,087	1,200
	(9.4%)	(90.6%)	(100%)

Table A.3. Region and Internet use (5 = extensive use, 1 = least use, 0 = no Internet access)

	5	4	3	2	1	0	Total
Luzon	286	78	53	10	8	165	600
	(47.7%)	(13.0%)	(8.8%)	(1.7%)	(1.3%)	(27.5%)	(100%)
Visayas	81	58	33	9	4	115	300
	(27.0%)	(19.3%)	(11.0%)	(3.0%)	(1.3%)	(38.3%)	(100%)
Mindanao	79	43	33	10	7	128	300
	(26.3%)	(14.3%)	(11.0%)	(3.3%)	(2.3%)	(42.7%)	(100%)
Total	446	179	119	29	19	408	1,200
	(37.2%)	(14.9%)	(9.9%)	(2.4%)	(1.6%)	(34.0%)	(100%)

Table A.4. Socioeconomic class and Internet use, face-to-face survey (5 = extensive use, 1 = least use, 0 = no Internet access)

	5	4	3	2	1	0	Total
ABC	63	15	7	1	3	24	113
	(55.8%)	(13.3%)	(6.2%)	(0.9%)	(2.7%)	(21.2%)	(100%)
DE	383	164	112	28	16	384	1,087
	(35.2%)	(15.1%)	(10.3%)	(2.6%)	(1.5%)	(35.3%)	(100%)
Total	446	179	119	29	19	408	1,200
	(37.2%)	(14.9%)	(9.9%)	(2.4%)	(1.6%)	(34.0%)	(100%)

Table A.5. Socioeconomic class and perceived neighborhood satisfaction with Duterte (4 = very satisfied, 1 = very dissatisfied)

	4	3	2	1	Total
ABC	21	61	15	5	102
	(20.6%)	(59.8%)	(14.7%)	(4.9%)	(100%)
DE	308	566	133	22	1,029
	(29.9%)	(55.0%)	(12.9%)	(2.1%)	(100%)
Total	329	627	148	27	1,131
	(29.1%)	(55.4%)	(13.1%)	(2.4%)	(100%)

Table A.6. Region and perceived neighborhood satisfaction with Duterte, face-to-face survey (4 = very satisfied, 1 = very dissatisfied)

	4	3	2	1	Total
Luzon	77	367	84	8	536
	(14.4%)	(68.5%)	(15.7%)	(1.5%)	(100%)
Visayas	101	137	41	18	297
	(34.0%)	(46.1%)	(13.8%)	(6.1%)	(100%)
Mindanao	151	123	23	1	298
	(50.7%)	(41.3%)	(7.7%)	(0.3%)	(100%)
Total	329	627	148	27	1,131
	(29.1%)	(55.4%)	(13.1%)	(2.4%)	(100%)

Table A.7. Region and perceived neighborhood satisfaction with Duterte, online survey (4 = very satisfied, 1 = very dissatisfied)

	4	3	2	1	Total
Luzon	559	918	453	175	2,105
	(26.6%)	(43.6%)	(21.5%)	(8.3%)	(100%)
Visayas	280	397	158	57	892
	(31.4%)	(44.5%)	(17.7%)	(6.4%)	(100%)
Mindanao	426	265	61	19	771
	(55.3%)	(34.4%)	(7.9%)	(2.5%)	(100%)
Total	1,265	1,580	672	251	3,768
	(33.6%)	(41.9%)	(17.8%)	(6.7%)	(100%)

C.3. Subgroup analyses by ethnicity

Dulay, Hicken and Holmes (2022) demonstrated that ethnicity was the most influential factor in explaining support for Duterte in the Philippines. Specifically, they showed that non-Tagalog people were much more likely to support Duterte than Tagalog people. However, their data were based on responses to direct questions, they expressed concern about the possibility that SDB confounded their results. That is, if non-Tagalog people were more susceptible to SDB, non-Tagalog people's position in favor of Duterte might be an artifact.

In response to that concern, we conducted a subgroup analysis by ethnicity, using the method employed for other subgroup analyses and reported in the "Explanatory subgroup analysis" section in the main text and Section C.1 in this file. Because our online survey lacked information about respondents' ethnicity, we only used the face-to-face survey. We used the variable of respondent ethnicity provided by Pulse Asia and categorized respondents into two groups: Tagalog people and non-Tagalog people.

Figure A.11 shows the results. The point estimate of SDB was slightly greater for non-Tagalog people, but the difference in SDB between Tagalog and non-Tagalog people was not statistically significant. Although we should refrain from interpreting statistical insignificance as evidence of no difference, we can at least conclude that we did not find signs indicating that the results reported by Dulay, Hicken and Holmes (2022) are an SDB-induced artifact.

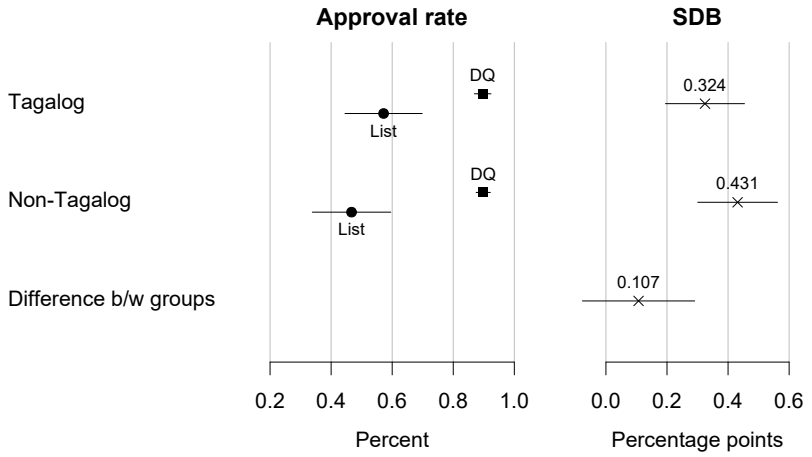


Figure A.11. Estimates of Duterte's approval ratings and SDB compared across ethnicity

Note: Cross marks and dots represent point estimates, and horizontal segments represent the 95% CIs.

References

Agerberg, Mattias and Marcus Tannenberg. 2021. "Dealing with Measurement Error in List Experiments: Choosing the Right Control List Design." *Research & Politics* 8(2): 1–8.

Ahlquist, John S. 2018. "List Experiment Design, Non-Strategic Respondent Error, and Item Count Technique Estimators." *Political Analysis* 26(1): 34–53.

Aronow, Peter M., Alexander Coppock, Forrest W. Crawford and Donald P. Green. 2015. "Combining List Experiment and Direct Question Estimates of Sensitive Behavior Prevalence." *Journal of Survey Statistics and Methodology* 3(1): 43–66.

Blair, Graeme and Kosuke Imai. 2010. "list: Statistical Methods for the Item Count Technique and List Experiment.". https://CRAN.R-project.org/package=list.

Blair, Graeme and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20(1): 47–77.

Dulay, Dean, Allen Hicken and Ronald Holmes. 2022. "The Persistence of Ethnopopulist Support: The Case of Rodrigo Duterte's Philippines." *Journal of East Asian Studies* 22(3): 525–553.

Kuhn, Patrick M. and Nick Vivyan. 2022. "The Misreporting Trade-Off between List Experiments and Direct Questions in Practice: Partition Validation Evidence from Two Countries." *Political Analysis* 30(3): 381–402.

Riambau, Guillem and Kai Ostwald. 2021. "Placebo Statements in List Experiments: Evidence from a Face-to-Face Survey in Singapore." *Political Science Research and Methods* 9(1): 172–179.