ABSTRACT

Large majorities in nearly every country support democracy, according to studies of cross-national surveys. But many of these reports have treated as missing data persons who did not provide a substantive response when asked to offer an opinion about the suitability of democracy as a regime type for their country, which has led to substantial overestimates of expressed support for democracy in some countries. This article discusses the consequences of excluding such nonsubstantive responses and offers suggestions to improve the study of popular support for democracy.

THE PROBLEM WITH EXCLUDING NONSUBSTANTIVE RESPONSES

Perhaps the most basic measure of popular support for democracy in a country is the percent of the population that agrees that democracy would be a good political system for their country. The English version of a World Values Survey item measuring support for democracy along these lines reads:

I’m going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad, or very bad way of governing this country? ... Having a democratic political system.

Figure 1 presents results to this item from the 2,532 respondents to the 2000 Iran Values Survey: 1,397 respondents indicated that democracy would be a “very good” or “fairly good” way to govern Iran, represented by the black line from 0% to the black dot at 55%; 235 respondents indicated that democracy would be a “fairly bad” or “very bad” way to govern Iran, represented by the gray line from 91% to 100%; and 900 respondents did not indicate whether democracy would be a good or bad way to govern Iran, represented by the gap from 55% to 91%.

Therefore, 55% of Iranian respondents at least paid lip service to the idea that democracy would be a good way to govern their country. But this percentage is artificially inflated when, as is commonly done, the calculation of support excludes observations of persons who did not indicate whether democracy would be a good or bad way to govern their country: if the 36% of the Iranian sample who did not select one of the “good way” or “bad way” options is excluded, then the 55% that expressed support is divided by the 64% that expressed support or opposition, resulting in a calculated level of 86% expressed support for democracy, represented by the gray dot in figure 1—an estimate that overstates expressed support by a full 31 percentage points.

But the 36% of Iranian respondents who declined to support or oppose democracy for their country should not be excluded: although these nonsubstantive responders have not rejected democracy with the same intensity as the 9% who expressly indicated that democracy would be a bad way to govern Iran, both groups have declined to endorse democracy as a good political system for their country. The practice of excluding nonsubstantive responses when reporting expressed support presumes that a situation of 86% expressed support, 14% expressed opposition, and 0% nonsubstantive response is equivalent to a situation of 55% expressed support, 9% expressed opposition, and 36% nonsubstantive response, and—for that matter—that both of these situations would be equivalent to a situation of 6% expressed support, 1%
expressed opposition, and 93% nonsubstantive response. In each case, 86% of substantive responders support democracy, but a sample in which 86% express support for democracy is not equivalent to a sample in which 6% express support for democracy.

Figure 2 recreates figure 1 for the countries listed in table 1 of Inglehart (2003), with responses from East and West Germany combined. Gray dots represent the percentage expressed support for democracy reported in the article, but the percentage that includes nonsubstantive responses is indicated by the black dot at the end of the black line.4 The gap between these percentages is often small, but the gap tends to be large in the nondemocratic countries for which measurement of support for democracy is critical.3 For example, in 2001, 68% of respondents in Morocco and 73% of respondents in China indicated that democracy would be a good way to govern their country, but expressed support for democracy in these countries had been reported as 96% in studies that excluded nonsubstantive responses (see Dalton and Shin 2006, 82; Inglehart 2003, 52; Inglehart and Norris 2003, 66; and Tessler 2002, 236).

Excluding nonsubstantive responses can also affect inferences about the correlates of support for democracy, given that nonsubstantive response can be associated with such predictors. For example, in the 2000 Iran Values Survey, nonsubstantive response to the “having a democratic political system” item negatively correlated with higher levels of formal education ($r = -0.23, n = 1,632, p < 0.01$), positively correlated with increased age ($r = 0.09, n = 1,632, p < 0.01$), and was higher among women than among men (41% and 31%, respectively, $n = 1,632$).

To illustrate the variation in inferences due to exclusion of nonsubstantive responses, we calculated predicted probabilities with CLARIFY (King, Tomz, and Wittenberg 2000; Tomz, Wittenberg, and King 2003) based on two logistic regressions using gender, age, and education as predictor variables; the dichotomous dependent variable in both regressions was coded 1 if the respondent indicated that democracy was a good way to govern Iran, and 0 if the respondent did not indicate that democracy was a good way to govern Iran. The first regression excluded nonsubstantive responders, and the second regression included nonsubstantive responders in the 0 category. This coding decision altered inferences about the influence of education on support for democracy and about relative levels of support for democracy: when nonsubstantive responses were excluded, with gender and age at their mean values, 95% confidence intervals for the predicted probabil-

SUGGESTIONS FOR HANDLING NONSUBSTANTIVE SURVEY RESPONSES

Studies should include nonsubstantive responses when calculating expressed support and should report the percentage of nonsubstantive responses so that readers see the range of possible levels of true support. For example, expressed support for a democratic Iran based on the 2000 Iran Values Survey sample could be presented as a [55, 91] percent interval, indicating that 55% of respondents expressed support, 9% of respondents expressed opposition, and the 36% reflected in the 55-to-91 interval did not provide a substantive response. This provides more and better information than an 86% point estimate and permits the reader to surmise that the level of true support for a democratic Iran in the sample is between 55% and 91%, depending on the percentage of nonsubstantive responders who censored their support for a democratic Iran.5

But studies concerned with providing a more precise estimate of true support for democracy should report a thorough analysis of nonsubstantive responses to assess as best as possible the percentage of nonsubstantive responses that reflect censored support or censored opposition. For this task, omission of nonsubstantive responses—a practice known as listwise deletion—is not necessarily a bad technique: listwise deletion might inflate reports of expressed opinion (what respondents said), but listwise deletion does not bias reports of true opinion (what respondents think) as long as the distribution of opinion among substantive responders matches the distribution of opinion among nonsubstantive responders. For example, 86% as an estimate of true support for a democratic Iran is accurate to the extent that the ratio of true support among the 900 nonsubstantive responders matches the 86% expressed—and presumably true—support among the 1,632 substantive responders.

Listwise deletion is equivalent to treating observed cases as a random sample of all cases and thus assuming that unobserved cases are missing completely at random (see Allison 2002, 646). But this is often not the case. In the 2000 Iran Values Survey
sample, for example, women were more likely than men to not provide a substantive response to the “having a democratic political system” item. For this situation, in which nonsubstantive responses are not missing completely at random, researchers should use techniques that soften the missing-completely-at-random assumption, such as inverse probability weighting, multiple imputation, and Heckman selection modeling.

Inverse Probability Weighting

In inverse probability weighting, observed data are used to predict the probability of providing a substantive response; the value of the substantive response from each observed respondent is then weighted by the inverse of the predicted probability that a respondent with his or her characteristics provided a substantive response.

Our first inverse probability weighting analysis modeled the presence of a substantive response to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey. This model used a regression equation with a variable for gender, a variable for age, and a formal education variable that ranged from 1 for incomplete elementary education to 8 for a university degree. Resulting predicted probabilities of a substantive response from a given respondent ranged from 39% to 85%; for example, a 65-year-old woman with an incomplete elementary education had a 43% chance of providing a substantive response. The inverse values of these probabilities were then used to weight substantive responses. For example, substantive responses from 65-year-old women with an incomplete elementary education were weighted by a factor of 2.3, which is the inverse of 43% (i.e., 1 divided by 0.43). Inverse probability weights ranged from 1.2 to 2.6, with a respective mean and standard deviation of 1.6 and 0.3. This first inverse probability weighting analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

Our second inverse probability weighting analysis modeled the presence of a substantive response to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey using variables for gender, age, and formal education, plus variables for opinion about a political system with a strong leader, opinion about a political system in which experts make decisions, opinion about a political system in which the army rules, and agreement with the statement that democracy may have its problems but it is better than any other political system. Responses for these latter four opinion measures were trichotomized into variables in which 1 indicated a response of good or agreement, −1 indicated a response of bad or disagreement, and 0 indicated a nonsubstantive response.
Inverse probability weights ranged from 1.1 to 5.5, with a respective mean and standard deviation of 1.7 and 0.6. This second inverse probability weighting analysis produced a 95% confidence interval of [81, 86] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

**Multiple Imputation**

In multiple imputation, observed data predict multiple values of a missing value for an observation. The range of variation in predicted values reflects the uncertainty about the value that should be predicted; then, these multiple predicted values are used to generate multiple complete datasets, such that n predicted values produce n complete datasets. These n complete datasets are analyzed, and the results of the analyses are combined with statistical rules.

Our first multiple imputation used gender, age, and education to predict responses to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey who did not provide a response to that item; three responses were predicted for each missing value, to generate three complete datasets. This first multiple imputation analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

Our first multiple imputation used gender, age, and education to predict responses to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey who did not provide a response to that item; three responses were predicted for each missing value, to generate three complete datasets. This first multiple imputation analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

The third multiple imputation was based on the aforementioned seven variables, plus variables for confidence in government, a 1-to-10 rating of the current political system for the country, opinion about whether the economic system runs badly in a democracy, opinion about whether democracies are indecisive, opinion about whether democracies are not good at maintaining order, satisfaction with the financial situation of the household, and frequency of attendance at religious services. This third multiple imputation analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

**Heckman Selection Models**

Multiple imputation and inverse probability weighting presume that the probability of a nonsubstantive response is not a function of the response variable. This assumption cannot be tested statistically and is unlikely to hold in the nondemocratic regimes where measurement of support for democracy is critical. In these regimes, nonsubstantive response due to self-censorship might be more likely among persons who prefer replacing the current regime with a democratic regime than among persons who prefer the nondemocratic status quo. This assumption is especially problematic considering that, in our cross-national analysis, the nonsubstantive response rate to democratic support items was inversely related to the level of democracy in a country.

Heckman selection models can account for missing data when the probability of nonsubstantive response is a function of the response variable: observed data predict the probability of providing a substantive response, and a transformation of this probability is included as a control in a regression to predict the dependent variable of interest (Heckman 1976). But if the selection model is not specified correctly, a Heckman model can produce estimates inferior to estimates from models assuming that nonsubstantive response is not a function of the response variable (Enders 2010, 296).

Our first Heckman model predicted the presence of a substantive response to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey using the seven variables from the second inverse probability regression equation plus a self-reported interest-in-politics variable. The first seven variables were used as predictors in the outcome equation; the regres-

This provides more and better information than an 86% point estimate and permits the reader to surmise that the level of true support for a democratic Iran in the sample is between 55% and 91%, depending on the percentage of nonsubstantive responders who censored their support for a democratic Iran.

Our second multiple imputation used the seven variables from the second inverse probability regression equation to predict responses to the “having a democratic political system” item among respondents to the 2000 Iran Values Survey who did not provide a response to that item; three responses were predicted for each missing value, to generate three complete datasets. This second multiple imputation analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

The third multiple imputation was based on the aforementioned seven variables, plus variables for confidence in government, a 1-to-10 rating of the current political system for the country, opinion about whether the economic system runs badly in a democracy, opinion about whether democracies are indecisive, opinion about whether democracies are not good at maintaining order, satisfaction with the financial situation of the household, and frequency of attendance at religious services. This third multiple imputation analysis produced a 95% confidence interval of [84, 87] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

Our second Heckman model retained the variables from the first Heckman model and added outcome predictors of financial satisfaction, confidence in the government, confidence in the churches, and frequency of religious service attendance. These variables were also included as predictors in the selection equation, alongside new variables for self-reported importance of politics to the respondent and for the frequency of political discussions with friends. Because data were missing in these additional variables, the regression sample size was reduced to 1,156 uncensored observations and 490 censored observations. This second Heckman analysis produced a 95% confidence interval of [76, 91] percent true support for a democratic Iran among respondents to the 2000 Iran Values Survey.

**Summary**

Given the diversity of techniques available to estimate censored opinions—and the diversity of assumptions that they require—reports of estimated true support for democracy should reflect...
multiple methods for handling censored responses to inform readers about the sensitivity of estimates to missing data techniques and assumptions. Estimates of true support for democracy in the 2000 Iran Values Survey provide an example of relative convergence: the 86% estimate of true support for democracy derived from listwise deletion fell close to or within the 95% confidence intervals derived from inverse probability weighting, multiple imputation, and Heckman selection modeling.

**IMPROVING THE DESIGN OF DEMOCRATIC SUPPORT ITEMS TO REDUCE NONSUBSTANTIVE RESPONSE**

Given the lack of an optimal statistical adjustment for reducing uncertainty caused by nonsubstantive response, expectations about even the most careful posthoc adjustment should be tempered by the multitude of possible reasons for nonsubstantive response to items that measure support for democracy, such as censorship of prodemocracy views considered at odds with the regime party line, censorship of antidemocracy views considered politically or socially undesirable, a general reluctance to express opinions stemming from limited formal education or other characteristics, and a lack of clarity about the actual nature of a democratic system (see Chu and Huang 2010).

Thus, the best way to reduce nonsubstantive response levels might be to modify the items themselves. For example, collapsing the four “having a democratic political system” response options into a dichotomous choice might increase the substantive response rate by decreasing task difficulty; respondents would need to map their opinion onto only a single directional dimension—good or bad—instead of both a directional dimension and an intensity dimension—very good, fairly good, fairly bad, or very bad. In addition, substantive response rates might be increased by offering a neutral option to capture opinions located between good and bad, because lack of a middle option can induce nonresponse (see Krosnick and Presser 2010, 271, 274). Neutrality or ambivalence about democracy should be captured because it is a legitimate attitude, especially among respondents in countries who lack democratic experience. For example, 77% of respondents to the 2000 Iran Values Survey who indicated that democracy would be a good way to govern Iran also endorsed military rule as a good way to govern Iran.

**IMPLICATIONS FOR THE STATUS OF GLOBAL SUPPORT FOR DEMOCRACY**

Because support for democracy might be a prerequisite for democratization, proper handling of nonsubstantive responses to democratic support items can have implications for democracy promotion efforts that rely on public opinion surveys.

### Notes

1. See, for example, Dalton and Shin 2006, 95; Esmer 2009, 291; Inglehart 2003, 52; Inglehart and Norris 2003, 52; Mishler and Rose 2003, 516; and Fletcher 2002, 236. Other studies, however, include nonsubstantive responses as valid observations: see, for example, Carballo 2008, 97; Dalton, Shin, and Jou 2007, 146; Kotez 2008, 357; Moreno and Méndez 2003, 234; Rose 2002, 107; and Shi and Lu 2010, 128.

2. Responses to the democratic support item were described with terms such as "say" and "endorse," indicating that reported percentages were intended to reflect expressed support.

3. There is evidence that the probability of a nonsubstantive response to democratic support items is a function of the response variable: country-level nonsubstantive response rate to democratic support items negatively correlated with the country’s democracy rating. The correlation is $-0.30 (p = 0.01, n = 73)$ for the “having a democratic political system” item, based on the respective European Values Study or World Values Survey for countries shown in figure 2. The correlation is $-0.43 (p = 0.01, n = 70)$ for countries whose respondents were asked to agree or disagree with the statement that “Democracy may have problems but it’s better than any other form of government.” Democracy ratings were drawn from the politva variable in the Marshall, Gurr, and Jaggers Polity IV Annual Time-Series 1800–2009 dataset, p4v2009.xls, retrieved February 1, 2011, from http://www.systemicpeace.org/inscr/inscr.htm. Ratings for Bosnia and Herzegovina and for Serbia were drawn from the Polity IV Individual Country Regime Trends 1946–2008 graphs available at http://www.systemicpeace.org/polity/polity4.htm; ratings were not available for Iceland, Luxembourg, and Malta; the rating for the United States was used for Puerto Rico, and the rating for the United Kingdom was used for Great Britain and for Northern Ireland.

4. The 95% confidence interval for the predicted probability of support for democracy still ranged from [33, 41] percent for the lowest education level to [59, 75] percent for the highest education level when a multinomial logistic regression predicted a three-value dependent variable with separate categories for responses of “good way,” “bad way,” and nonselection of “good way” or “bad way.” The multinomial regression was used instead of an ordered logistic regression because a Brant test revealed that all three model independent variables violated the parallel lines assumption of the ordinal logistic regression when the trichotomous dependent variable ranged from a response of “good way” to a nonsubstantive response to a response of “bad way.”

5. This confidence interval was constructed based solely on uncertainty due to nonsubstantive response, to illustrate the degree and variability of estimate uncertainty caused by nonsubstantive response. But, when possible, confidence intervals can also account for other survey characteristics, such as sample size, total refusal rate, and sampling error.

6. Standard errors for this and other inverse probability weighting estimates were not adjusted to account for the percent of nonsubstantive responses. Confidence intervals for this and other estimates of true support were not adjusted to reflect sampling error.
7. The imputation was conducted with the Stata 11 mi routine, with the command:
\texttt{mi impute mvn [dependent variable] = [independent variables], replace add(3) rseed(123) force.}

\textbf{REFERENCES}


Robert Wood Johnson Foundation
Scholars in Health Policy Research

Congratulates Political Science Recipients of the 2012–2014 Fellowship Awards

Daniel Gillion
(Harvard University Site)
Assistant Professor,
University of Pennsylvania
2009 PhD, University of Rochester

Andrew Kelly
(University of California, Berkeley/San Francisco Site)
2012 PhD, Northwestern University

Francisco Pedraza
(University of Michigan Site)
Assistant Professor,
Texas A&M University
2010 PhD, University of Washington

Robert Wood Johnson Foundation Scholars in Health Policy Research is a national fellowship program of the Robert Wood Johnson Foundation. Each year, it provides up to 9 talented social scientists with a unique and challenging two-year post-doctoral research and training experience at one of three participating universities: The University of California at Berkeley and San Francisco; The University of Michigan; and Harvard University.

For information about the program, please visit www.healthpolicyscholars.org.

CAMBRIDGE

Coming this Fall from Cambridge University Press

Making Democratic Governance Work
How Regimes Shape Prosperity, Welfare, and Peace
Pippa Norris
$90.00: Hb: 978-1-107-01699-6
$29.99: Pb: 978-1-107-60269-4
288 pp.

Democracy Prevention
The Politics of the U.S.-Egyptian Alliance
Jason Brownlee
$85.00: Hb: 978-1-107-02571-4
312 pp.
Prices subject to change.

Constitutional Change and Democracy in Indonesia
Donald L. Horowitz
Problems of International Politics
$90.00: Hb: 978-1-107-02727-5
$24.99: Pb: 978-1-107-64115-0
328 pp.

The Politics of Authoritarian Rule
Milan W. Svolik
Cambridge Studies in Comparative Politics
$85.00: Hb: 978-1-107-02479-3
$29.99: Pb: 978-1-107-60745-3
288 pp.

Evaluating Elections
A Handbook of Methods and Standards
R. Michael Alvarez, Lonna Rae Atkeson, and Thad Hall
$85.00: Hb: 978-1-107-02762-6
$27.99: Pb: 978-1-107-65305-4
200 pp.

Natural Experiments in the Social Sciences
A Design-Based Approach
Thad Dunning
Strategies for Social Inquiry
$99.00: Hb: 978-1-107-01766-5
$36.99: Pb: 978-1-107-69800-0
368 pp.

www.cambridge.org/us