

Regressing the Propensity to Vote

Gordon G. Bechtel

University of Florida and Florida Research Institute

Abstract: The present paper addresses the propensity to vote with data from the third and fourth rounds of the European Social Survey. The regression of voting propensities on true predictor scores is made possible by estimates of predictor reliabilities (Bechtel, 2010; 2011). This resolves two major problems in binary regression, i.e. errors in variables and imputation errors. These resolutions are attained by a pure randomization theory that incorporates fixed measurement error in design-based regression. This type of weighted regression has long been preferred by statistical agencies and polling organizations for sampling large populations.

Key words: Binary propensity, coefficient alpha, interval scales, measurement error, ordinary least squares, randomization theory, true-value regression.

1. Introduction

The present paper treats the propensity to vote (or not vote) in a national election as taking non-binary continuous values on an interval scale. This propensity is regressed on the true values of explanatory variables that have been cleansed of error (Bechtel, 2010; 2011). First, an individual's reported voting, not voting, or imputation (in the case of item non-response) is interpreted as an interval-scale propensity plus a measurement error. Second, each of her (his) predictor scores is also interpreted as a true interval-scale value plus an error score. Then, classical assumptions about measurement error (Gulliksen, 1950, pp. 4-7; Bound, Brown and Mathiowetz, 2001), along with reliability coefficient alpha in psychological test theory (Cronbach, 1951; Lord and Novick, 1968; Nunnally and Bernstein, 1994; StataCorp., 2001), allow the regression of voting propensities on true predictor values.

This treatment of "errors in variables" avoids likelihood maximization and its assumption that a census of survey scores is itself a sample from a "superpopulation" with a specified distribution. (Skinner, Holt and Smith, 1989; Valliant,

Dorfman and Royall, 1999). The present approach also avoids the further model-based assumptions that a) true predictor scores are normally distributed and b) measurement-error variance is known (Fuller, 1987, pp. 103-106; Bound, Brown and Mathiowetz, 2001). Here, no distribution assumptions are made about true predictor scores, and measurement-error variance is obtained from estimated reliabilities of predictor scores. This allows measurement error to be circumvented by the estimation of census totals.

The present more parsimonious approach to “errors in variables” adds measurement error to design-based regression, thus extending Neyman (1934) randomization theory (Bellhouse, 1988; Nathan, 1988; Thompson, 1997; StataCorp., 2001, Volume 4, pp. 29-30; Chaudhuri and Stenger, 2005; Lohr, 2010, pp. 434-443). This extension is accomplished by distinguishing a true-value population from a realized census of binary responses and observed predictors. Sections 2 and 3 describe binary voting responses, their predictors, and measurement errors in these variables. Section 4 defines our population of voting propensities, true predictor values, and the target parameter of an OLS regression over this population. Section 5 develops an estimator for this target using weighted sample totals that estimate corresponding census totals. Section 6 demonstrates this estimation with cross-national datasets from the third and fourth rounds of the European Social Survey (ESS). Section 7 points up true-value theory as a “pure” extension of randomization theory that realistically addresses micro data in public opinion polling.

2. Responses, Imputations, and Propensities

To measure voting propensity we use the ESS question on voting, which is phrased as follows:

Some people don't vote nowadays for one reason or another. Did you vote in the last [country] national election in [month/year]? *No* 0 *Yes* 10 (<http://ess.nsd.uib.no>)

This question refers to the last election of a country's primary legislative assembly. Individual i 's response is coded zero or ten, whereas a missing response is filled in as an imputation that usually lies between these two values. Letting Y_i be a response or imputation, we model it as

$$Y_i = \eta_i + E_i. \quad (2.1)$$

In (2.1) η_i is i 's propensity to vote and E_i is i 's response or imputation error. The propensity and error on the right side of (2.1) lie on a continuous interval scale whose origin and unit are set by the coding of the response labels *no* and *yes*. Thus we regard the departure of zero or ten from η_i as measurement error. We

also interpret the departure of an imputation from this interval-scale propensity as measurement error.

3. Predictor Scores and Their True Values

The present study regresses voting propensity on the two multi-item predictors in Table 1, along with party proximity and age, which are single-item predictors. The ESS question on party proximity is phrased and coded as follows:

Is there a particular political party you feel closer to than all the other parties? *No* 0
Yes 10 (<http://ess.nsd.uib.no>)

Table 1: Multi-item scales for predicting voting propensity

Political efficacy scale

How interested would you say you are in politics - are you...

not at all	0	hardly	3.33	quite	6.67	very interested	10
------------	---	--------	------	-------	------	-----------------	----

How often does politics seem so complicated that you can't really understand what is going on?

frequently	0	regularly	2.5	occasionally	5.0	seldom	7.5	never	10
------------	---	-----------	-----	--------------	-----	--------	-----	-------	----

How difficult or easy do you find it to make your mind up about political issues?

very difficult	0	difficult	2.5	neither difficult nor easy	5.0	easy	7.5	very easy	10
----------------	---	-----------	-----	----------------------------	-----	------	-----	-----------	----

Political trust scale

How much do you personally trust each of the following institutions. [country]'s politicians; political parties?

no trust at all	0	1	2	3	4	5	6	7	8	9	10	complete trust.
-----------------	---	---	---	---	---	---	---	---	---	---	----	-----------------

Source: These items are found on the website ess.nsd.uib.no

The five item responses used to measure political trust and efficacy are coded in Table 1. Missing responses are filled in as imputations that lie among these coded values. An individual's score on each of these multi-item predictors is the average of her (his) item ratings/imputations and ranges from 0 to 10. In order to compare regression coefficients, we have calibrated party proximity and age to range between 0 and 10 as well.

The composite scores for political trust and efficacy in Table 1 contain measurement error. The single-item scores for party proximity and age are assumed to be error free.

3.1 Errors in Political Efficacy Scores

Referring to the first scale in Table 1, we write respondent i 's three item

ratings/imputations as

$$\begin{aligned} X_{i11} &= \tau_{i1} + U_{i11}, & (\text{interest}), \\ X_{i12} &= \tau_{i1} + U_{i12}, & (\text{understanding}), \\ X_{i13} &= \tau_{i1} + U_{i13}, & (\text{issue resolution}), \end{aligned}$$

where her (his) political efficacy τ_{i1} lies on a continuous interval scale. The origin and unit of this scale are set by coding the extreme response labels as zero and ten for each of these items. The departure U_{i1m} for $m = 1, 2, 3$ is a coded-response or imputation error in measuring τ_{i1} with item m . Thus, if U_{i11} is a response error, it is a departure of the coded value 0, 3.33, 6.67 or 10 from τ_{i1} on our interval political-efficacy scale. If U_{i12} is a response error, it is a deviation of 0, 2.5, 5.0, 7.5 or 10 from τ_{i1} on this same scale. This interval scale tolerates items with differing numbers of response options. It also tolerates the equal spacing of response options that is ubiquitously used in survey work. If this spacing is subjectively incorrect for respondent i , then U_{i1m} is an increased coding error.

Finally, averaging over our three item scores gives individual i 's political efficacy score as

$$\begin{aligned} X_{i1} &= (X_{i11} + X_{i12} + X_{i13})/3 \\ &= \tau_{i1} + (U_{i11} + U_{i12} + U_{i13})/3 \\ &= \tau_{i1} + U_{i1}. \end{aligned} \tag{3.1}$$

The U_{i1} in (3.1) is individual i 's error score, which can be a mixture of item response and imputation errors.

3.2 Errors in Political Trust Scores

Still referring to Table 1, we write respondent i 's ratings/imputations of the political trust items as

$$\begin{aligned} X_{i21} &= \tau_{i2} + U_{i21}, & (\text{politicians}), \\ X_{i22} &= \tau_{i2} + U_{i22}, & (\text{political parties}), \end{aligned}$$

where τ_{i2} is the interval-scale value of individual i 's political trust. The departure U_{i2m} for $m = 1, 2$ is an item error in measuring τ_{i2} . Again we regard the departure of the coded value X_{i2m} from τ_{i2} as measurement error. We also interpret the departure of an imputation from τ_{i2} as measurement error. Individual i 's political trust score is

$$\begin{aligned} X_{i2} &= (X_{i21} + X_{i22})/2 \\ &= \tau_{i2} + (U_{i21} + U_{i22})/2 \\ &= \tau_{i2} + U_{i2}, \end{aligned} \tag{3.2}$$

where the error score U_{i2} is the average error in measuring i 's true political trust τ_{i2} .

4. Population Inference via an Imputed Census

4.1 A Generalization of Randomization Theory

We now posit a hypothetical (but possible) census which establishes a link to our population of true values (cf. Bechtel, 2010; 2011). We assume that missing item responses in this census have been imputed in the same manner as the sample imputations described in Section 6.3. In this imputed census the voting value Y_i in (2.1), along with the efficacy and trust scores X_{i1} and X_{i2} in (3.1) and (3.2), are constants. The errorless political-proximity and age responses, $X_{i3} = \tau_{i3}$ and $X_{i4} = \tau_{i4}$, are also census constants for the i -th individual. These fixed (rather than random) census values are in keeping with the fact that a census is only conducted once. It follows that each measurement error in Sections 2 and 3 is fixed because it is the difference between a fixed census value and a fixed true value for individual i . Thus, our population, error set, and census are three finite sets of constants that are in one-to-one correspondence. These sets are

$$\begin{aligned} & \{\eta_i, \tau_{i1}, \tau_{i2}, \tau_{i3}, \tau_{i4} \mid i = 1, \dots, N\}, \\ & \{E_i, U_{i1}, U_{i2}, 0, 0 \mid i = 1, \dots, N\}, \text{ and} \\ & \{Y_i, X_{i1}, X_{i2}, X_{i3}, X_{i4} \mid i = 1, \dots, N\}, \end{aligned}$$

where N is the aggregate population size of the countries listed in Section 6.1. Individual i 's (one-time) census score for each variable is the sum of her (his) fixed error and true value. This incorporation of constant measurement error in design-based regression preserves a pure randomization theory in which sample inclusion (or not) for each individual $i = 1, \dots, N$ is her (his) only random variable.

4.2 Parameter Identification

Using $\{\eta_i, \tau_{i1}, \tau_{i2}, \tau_{i3}, \tau_{i4} \mid i = 1, \dots, N\}$, our population model is

$$\eta_i = \beta_0 + \beta_1\tau_{i1} + \beta_2\tau_{i2} + \beta_3\tau_{i3} + \beta_4\tau_{i4} + \varepsilon_i, \quad \text{for } i = 1, \dots, N, \quad (4.1)$$

where ε_i is a fixed specification error. These errors, along with the coefficients $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$, are uniquely identified by the ordinary-least-squares condition that $\sum \varepsilon_i^2$ is minimal when the propensities η_i are regressed on the true predictors

$\tau_{i1}, \tau_{i2}, \tau_{i3}, \tau_{i4}$ over $i = 1, \dots, N$. This OLS identification of the true slope vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ is

$$\boldsymbol{\beta} = \left[\sum (\boldsymbol{\tau}_i - \boldsymbol{\tau}_\cdot)(\boldsymbol{\tau}_i - \boldsymbol{\tau}_\cdot)^T \right]^{-1} \sum (\boldsymbol{\tau}_i - \boldsymbol{\tau}_\cdot)(\eta_i - \bar{\eta}), \quad (4.2)$$

where $\boldsymbol{\tau}_i = (\tau_{i1}, \tau_{i2}, \tau_{i3}, \tau_{i4})^T$, $\boldsymbol{\tau}_\cdot = (\tau_{\cdot 1}, \tau_{\cdot 2}, \tau_{\cdot 3}, \tau_{\cdot 4})^T$, and the two population summations run over $i = 1, \dots, N$.

4.3 An Estimable Form of the Population Target

The classical error assumptions. We now assume that over our census error set $\{E_i, U_{i1}, U_{i2}, 0, 0 \mid i = 1, \dots, N\}$ the error scores in (2.1), (3.1) and (3.2) sum to zero, are uncorrelated with true scores, and are uncorrelated with each other (Gulliksen, 1950, pp. 4-7; Bound, Brown and Mathiowetz, 2001). These classical assumptions may be written as the vanishing totals

$$\begin{aligned} \sum E_i &= \sum E_i \tau_{ij} = \sum E_i U_{ij} = 0, \quad \text{and} \\ \sum U_{ij} &= \sum U_{ij} \tau_{ij} = \sum U_{ij} \tau_{ik} = \sum U_{ij} U_{ik} = \sum U_{ij} \eta_i = 0, \end{aligned} \quad (4.3)$$

where the summations run over $i = 1, \dots, N$, and the predictors $j, k = 1, 2, 3, 4$.

Rewriting $\boldsymbol{\beta}$. Under the error assumptions in (4.3), along with (2.1), (3.1) and (3.2), it is easily shown that

$$\begin{aligned} \sum \tau_{ij} \eta_i &= \sum X_{ij} Y_i, \\ \sum \tau_{ij} \tau_{ik} &= \sum X_{ij} X_{ik}, \quad \text{for } j \neq k, \quad \text{and} \\ \sum \tau_{ij}^2 &= \sum X_{ij}^2 - \sum U_{ij}^2, \quad \text{for } j = k. \end{aligned}$$

Our population target (4.2) may then be written in the estimable form

$$\boldsymbol{\beta} = \left[\sum (\mathbf{X}_i - \mathbf{X}_\cdot)(\mathbf{X}_i - \mathbf{X}_\cdot)^T - \Delta \right]^{-1} \sum (\mathbf{X}_i - \mathbf{X}_\cdot)(Y_i - \bar{Y}), \quad (4.4)$$

where $\mathbf{X}_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4})^T$ and $\mathbf{X}_\cdot = (X_{\cdot 1}, X_{\cdot 2}, X_{\cdot 3}, X_{\cdot 4})^T$. Again the two population summations run over $i = 1, \dots, N$. In (4.4) the matrix $\Delta = \text{diag}(\delta_1, \delta_2, 0, 0)$, and

$$\delta_j = \sum U_{ij}^2 = (1 - \alpha_j) \left\{ \sum X_{ij}^2 - (\sum X_{ij})^2 / N \right\}, \quad (4.5)$$

for $j = 1$ (efficacy), 2 (trust). The diagonals $\delta_3 = \delta_4 = 0$ because party proximity $X_{i3} = \tau_{i3}$ and age $X_{i4} = \tau_{i4}$ are error-free variables. The α_1 and α_2 in (4.5) are

census reliability coefficients for our efficacy and trust scores (Cronbach, 1951; Lord and Novick, 1968; Nunnally and Bernstein, 1994; StataCorp., 2001). These alpha coefficients provide the census error sums of squares δ_1 and δ_2 in (4.5) (Bechtel, 2010).

The sufficiency of one census. Assume a distinct error set $\{\ddot{E}_i, \ddot{U}_{i1}, \ddot{U}_{i2}, 0, 0 \mid i = 1, \dots, N\}$ that satisfies (4.3) for any other (simultaneous) census $\{\ddot{Y}_i, \ddot{X}_{i1}, \ddot{X}_{i2}, \ddot{X}_{i3}, \ddot{X}_{i4} \mid i = 1, \dots, N\}$. Then the difference set is again our target population $\{\eta_i, \tau_{i1}, \tau_{i2}, \tau_{i3}, \tau_{i4} \mid i = 1, \dots, N\}$, and the new census values, entered in (4.4), again generate β in (4.2). Hence, there is no need to entertain a second census or a super-population of censuses. Nor is it necessary to regard measurement error as a random variable that takes values (within each individual) over census realizations from this super-population. Thus, our fixed error scores in (2.1), (3.1), and (3.2), with properties (4.3), keep true-value regression as a pure randomization theory.

5. Slope Estimates Corrected for Measurement Error

Let $\{Y_i, X_{i1}, X_{i2}, X_{i3}, X_{i4} \mid i = 1, \dots, n\}$ be a sample drawn from the census $\{Y_i, X_{i1}, X_{i2}, X_{i3}, X_{i4} \mid i = 1, \dots, N\}$, where n is the net ESS sample size over all the countries listed in Section 6.1. This sample provides the following Horvitz-Thompson type estimator of $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ in (4.2) and (4.4):

$$B = \left[\sum w_i (\mathbf{X}_i - \mathbf{X})(\mathbf{X}_i - \mathbf{X})^T - D \right]^{-1} \sum w_i (\mathbf{X}_i - \mathbf{X})(Y_i - Y). \quad (5.1)$$

The euroweight w_i in (5.1) is described in Section 6.2. This weight adjusts micro pan European data for each respondent’s sample inclusion probability, each country’s population size, and each country’s unit non-response (Bechtel, 2011). The two sample summations in (5.1) run over $i = 1, \dots, n$, and the matrix $D = \text{diag}(d_1, d_2, 0, 0)$, where

$$d_j = (1 - a_j) \left\{ \sum w_i X_{ij}^2 - \left(\sum w_i X_{ij} \right)^2 / \sum w_i \right\}, \quad (5.2)$$

for $j = 1$ (efficacy), 2 (trust). Again $d_3 = d_4 = 0$ because party proximity $X_{i3} = \tau_{i3}$ and age $X_{i4} = \tau_{i4}$ are without measurement error. The matrix D in (5.1) corrects the well-known design-based regression formula, which holds when $D = 0$ (Nathan, 1988, pp. 255-256; Thompson, 1997, pp. 106-107; StataCorp., 2001, Volume 4, pp. 29-30; Chaudhuri and Stenger, 2005, pp. 264-265; Lohr, 2010, pp. 435-442).

The computation of a_j in (5.2), which is an estimate of α_j in (4.5), is given by Bechtel (2010). The estimated alpha coefficients, a_1 and a_2 , for our political efficacy and trust scales are exhibited in Table 2 for 2006 and 2008.

The standard errors of the slopes B_1, B_2, B_3 and B_4 in \mathbf{B} in (5.1) are derived in the Appendix of Bechtel (2010). These standard errors, which appear in Table 2, contain the effects of measurement error on the variances of the slopes in our propensity regressions.

Table 2: Design-based regressions of voting propensity

	<i>True-value regression</i>		<i>Naïve regression</i>	
	<i>2006</i>	<i>2008</i>	<i>2006</i>	<i>2008</i>
<i>Alpha coefficient</i>				
Political efficacy	.68	.69	1.00	1.00
Political trust	.92	.92	1.00	1.00
<i>Regression slope</i>				
Political efficacy	.405 (.032)	.400 (.030)	.265 (.017)	.265 (.017)
Political trust	.123 (.019)	.133 (.018)	.128 (.017)	.138 (.016)
Party Proximity	.156 (.008)	.177 (.008)	.174 (.008)	.192 (.007)
Age	.412 (.018)*	.354 (.022)	.406 (.018)*	.350 (.022)
Sample size	36,088	37,174	36,088	37,174

Note: The alpha coefficients for the true-value regressions are observed. The alpha coefficients for the naive regressions are (inappropriately) assumed to be one, indicating perfect reliability. The regression slopes are weighted by the euroweights described in Section 6.2. Their standard errors are in parentheses. The star (*) indicates a difference in slopes between successive surveys that is significant beyond the .05 level. All other slope changes are not significant.

6. European Propensity to Vote

6.1 The Sample Data

The data for our analysis was supplied by the European Social Survey (Jowell and the Central Co-ordinating Team, 2008) which is

funded through the European Commission's fifth and sixth Framework Programme, the European Science Foundation and national funding bodies in each country

Data collection takes place every two years, by means of face to face interviews of around an hour in duration The questionnaire consists of a 'core' module lasting about half an hour which remains relatively constant from round to round . . . the core module aims . . . to monitor change and continuity in a wide range of socio-economic, socio-political, socio-psychological and socio-demographic variables. (www.europeansocialsurvey.org)

All survey items in the ESS have been standardized and appropriately translated from country to country. The items in the present study are from the ESS core module that was administered in 2006 and 2008. The voting item is described in Section 2, and its multiple-item predictors are exhibited in Table 1. These predictors are supported by political proximity and age in explaining voting propensity.

Our analysis includes all 21 countries that were surveyed in both the third and fourth rounds of the ESS; namely, Belgium, Bulgaria, Switzerland, Cypress, Germany, Denmark, Estonia, Spain, Finland, France, Great Britain, Hungary, Netherlands, Norway, Poland, Portugal, Russia, Sweden, Slovenia, Slovakia and Ukraine. In each country a representative probability sample was drawn from the residential population aged 15 and older according to the following specifications:

- The minimum effective sample size is 1,500 (or 800 for countries with less than 2 million inhabitants).
- The net sample size (number of realised interviews) is calculated as the product of the effective sample size and the design effect, i.e. $n_{\text{net}} = n_{\text{eff}} * DEFF = 1,500 * DEFF$. An estimate of the design effect DEFF for each country was provided by its sampling expert.
- The gross sample size is calculated as follows: $n_{\text{gross}} = n_{\text{net}} / (RR * ER)$, where RR (target is 70%) is the predicted response rate and ER is the eligibility rate. (www.europeansocialsurvey.org)

6.2 Weighting for Unit Non-Response

The ESS provides design weights and population size weights that should be used to construct weights representative of cross-national European populations. Design weights, which are normed inverses of sample inclusion probabilities, sum to each country's net sample size. A country's population size weight is

(population size aged 15 and over)/(net sample size in the data file * 10000).

Then

$$\text{euroweight} = (\text{design weight}) * (\text{population size weight})$$

insures that our weighted cross-national regressions represent each country in proportion to its population size. Bechtel (2011) shows that this euroweight is the normed ESS sampling weight that has undergone a weighting class adjustment for unit non-response. The weighting classes for this non-response adjustment are the 21 countries in each of our two pan-European surveys.

6.3 Imputation for Item Non-Response

Two separate regression imputations were carried out for our voting and party-proximity items. Each of these single-item variables was regressed on party

membership, gender and education to impute its missing responses. Three regression imputations were conducted for the three-item political efficacy scale in Table 1. Each item was regressed on the other two items making up this scale, as well as on party membership, gender and education. Finally, two regression imputations were carried out for political trust, with each item regressed on the other, party membership, gender and education (cf. StataCorp., 2001, Volume 2, pp. 69-71). These regression imputations avoided sample loss by preserving all of our 21-nation datasets from the third and fourth rounds of the ESS.

Each of these item imputations closely estimates its corresponding census imputation. For example, in (3.1) if individual i 's interest value X_{i11} is a sample imputation, it is a weighted sum of her (his) observed understanding, issue resolution, party membership, gender and education values. Due to our very large sample, the weights in this sum are extremely close to those from a census imputation. Because the census weights are applied to i 's same understanding, issue resolution, party membership, gender and education values, the census imputation of i 's political interest is also extremely close to our sample imputation X_{i11} . Hence, individual i 's political efficacy score X_{i1} in (3.1), which is the average of two item scores and one imputation, differs negligibly from her (his) census score on this construct. These score differences will (approximately) sum to zero over a large sample. Thus, our weighted totals using sample imputations in formulas (5.1) and (5.2) are almost identical to the weighted sample totals that would be obtained with census imputations for missing item responses.

6.4 True-Value Regression of Voting Propensity¹

Predictors. Our purpose here is to correct for unreliability in the two multi-item scales in Table 1 in order to compare the strengths of political trust and efficacy on the European propensity to vote. An American counterpart of the political trust scale has been used since 1958 by the National Election Studies at The University of Michigan. However, even recent work on the implications of political trust for voting behavior has not included its effect on voter turnout per se (Rudolph, 2005). In contrast, the National Election Studies have demonstrated that political efficacy, as measured by American items, bears a strong relationship to voter turnout (Mattei and Niemi, 2005).

Our single-item predictors are party proximity and age. Party proximity is defined in Section 3 as closeness (or not) to one particular party without reference to which party or partisan direction (e.g. liberal versus conservative). This appears to be a construct that is new to the political-science literature. Finally, it is well known that age is the strongest demographic predictor of voter turnout.

¹The Stata .do file and documentation for running true-value regression may be obtained by email from the author.

Its regression slope serves as a bench mark against which to compare the slopes of our three political predictors of voting propensity.

Procedure. The variance inflation factor for each predictor $j = 1, 2, 3, 4$ is $VIF_j = 1/(1 - R_j^2)$. R_j^2 is the squared multiple correlation coefficient when predictor j is regressed on the other three predictors (StataCorp., 2001, pp. 111-114). The four VIFF values for our predictors in 2006 and 2008 are all close to one. Thus there is no co-linearity problem among these predictors in the following survey regressions.

The first section of Table 2 reveals that each of our multiple-item predictors is measured with equal reliability in 2006 and 2008. The political trust score consistently shows higher reliability than the political efficacy score.

Using the alpha coefficients in Table 2, voting propensity was regressed on political efficacy, political trust, party proximity and age. This true-value regression was repeated over 2006 and 2008 for cross-national samples from the 21 countries listed in Section 6.1. The resulting regression slopes in the second section of Table 2 were computed from formulas (5.1) and (5.2). The number of cases for each of these regressions appears at the bottom of the table.

Findings. Due to the coding of all items on the same interval scale (cf. Sections 2 and 3), the four predictors in Table 2 may be compared (vertically) as to their effects on voting propensity. Our major finding is the differential strength of the efficacy and trust predictors. Political efficacy rivals age, which is a well-known and powerful predictor of voting propensity. This confirms the American National Election Studies results for political efficacy (Mattei and Niemi, 2005). Surprisingly, political trust, which is also a venerable and closely watched variable in the National Election Studies (Rudolph, 2005), is our weakest predictor of European voting propensity. It is slightly excelled by the new construct, party proximity.

Reading Table 2 horizontally, the effect of age dropped slightly, but significantly, between 2006 and 2008. However, the determinants of European voting propensity remain stable over “normal” (2006) and “stressful” (2008) times. If this consistent regression structure in Table 2 is confirmed in future European and American studies, it will inform attempts to improve western voter turnout. For example, increasing citizen efficacy through education would appear to be more important than campaigns attempting to elevate public trust.

6.5 Naïve Design-Based Regressions

Standard practice in regressing a binary variable is to compute design-based logistic coefficients, which are pseudo-maximum-likelihood estimates (StataCorp., 2001, Volume 4, pp. 30-31). However, logistic regression cannot handle missing

voting data, and it treats voting imputations as voting responses (StataCorp., 2001, Volume 2, p. 252). Moreover, there is a theoretical disadvantage of using logistic regression as a “control analysis” for our binary true-value regression in Table 2, i.e. the pseudo-likelihood itself is based on the census model

$$\theta_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + v_i, \quad \text{for } i = 1, \dots, N,$$

where v_i is an intra-individual random error. Individual i 's observed response is “positive” when her (his) unobserved random variable $\theta_i > 0$ and “negative” when $\theta_i < 0$. However, this likelihood (unrealistically) assumes v_i to be identically logistically distributed for all N individuals in the census.

A comparable control analysis for our propensity regression is provided by a naïve design-based regression which treats the predictors $\{X_{i1}, X_{i2}, X_{i3}, X_{i4} \mid i = 1, \dots, N\}$ in the census as a finite set of errorless constants. Thus, by setting $D = 0$ in (5.1) we have a classic design-based regression on the same imputed dataset.

The right side of Table 2 displays the naive regression slopes for 2006 and 2008. In contrast to the true slopes for political efficacy, the naïve coefficients and standard errors for this predictor are spuriously low in 2006 and 2008. This sharp attenuation is due to the lower reliabilities, also in Table 2, of our political efficacy scores. In contrast, the highly reliable political-trust scale, along with the errorless party proximity and age variables, exhibit similar coefficients and standard errors in our true-value and naïve regressions.

Table 2 shows that failure to correct for measurement error in an important predictor can drastically underestimate its influence on a response variable. In the present case this could mislead educational policy vis-à-vis the importance of improving political efficacy in the voting population.

7. Generalized Randomization Theory in Public Opinion Polling

The present paper compares regression slopes under a generalized randomization theory (Bechtel, 2010; 2011) with those under the classical randomization theory used by statistical agencies and polling organizations (Neyman, 1934; Horvitz and Thompson, 1952; Chaudhuri and Stenger, 2005; Lohr, 2010). This comparison, which uses the third and fourth rounds of the ESS, addresses measurement errors in political efficacy and trust scores that predict voting propensity.

Generalized randomization theory relaxes the strong assumption in design-based theory that a finite population is a set of errorless constants (Nathan, 1988; Lehmann, 1999; Chaudhuri and Stenger, 2005; Lohr, 2010). Here this classical population is replaced by two finite sets of vectors. The first set is a population of N true vectors, and the second is a census of N erroneous vectors. Each of

these two sets consists of real, rather than random, numbers. The interpretation of census values Y_i , X_{i1} and X_{i2} in (2.1), (3.1) and (3.2) as deviations from true interval-scale values η_i , τ_{i1} and τ_{i2} is a step forward in the Neyman paradigm (cf. Bellhouse, 1988). This extension of randomization theory, without invoking a specifically distributed super-population postulated by model-based theory (cf. Fuller, 1975; 1987), establishes measurement error within a pure design-based regression. This resolves two major issues in survey regression by a) allowing errors in variables and b) viewing imputation errors as special cases of measurement errors in these variables. The classical assumptions in (4.3) about the behavior of these errors over the census enable the regression of true response propensities on true predictor values.

The additive errors in (2.1), (3.1) and (3.2), as fixed deviations from true values, correct extreme interpretations of survey measures as errorless constants on the one hand or specifically distributed random variables on the other. This more realistic interpretation of micro-data is implemented in Table 2, which shows that bias is reduced by a weighted correction for measurement error in political efficacy scores. This type of correction should reduce estimation bias in other government surveys and opinion polls that use multiple-item scores to predict important survey variables.

Acknowledgements

The author thanks the JDS referee whose critique and suggestions have improved this work. He also appreciates the support given by the University of Florida's Warrington College of Business Administration. This article is a corollary and application of theoretical work appearing in Bechtel (2010). Our cross-national regressions were obtained by user-friendly access to the European Social Survey. Norwegian Social Science Data Services (NSD) is the data archive and distributor of the ESS data. None of the ideas or analyses here may be attributed to the European Social Survey or the University of Florida.

References

- Bechtel, G. G. (2010). True-value regression theory. *Journal of Data Science* **8**, 521-539.
- Bechtel, G. G. (2011). True-value regression with non-response. *Journal of Data Science* **9**, 501-512.
- Bellhouse, D. R. (1988). A brief history of random sampling methods. In *Handbook of Statistics*, Volume 6 (Sampling) (Edited by P. R. Krishnaiah and C. R. Rao), 1-14. North Holland, New York.

- Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of Econometrics*, Volume 5 (Edited by J. J. Heckman and E. Leamer), 3705-3843. Elsevier Science, Amsterdam.
- Chaudhuri, A. and Stenger, H. (2005). *Survey Sampling: Theory and Methods*, 2nd edition. Chapman and Hall / CRC Press, New York.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* **16**, 297-334.
- Fuller, W. A. (1975). Regression analysis for sample survey. *Sankhyā, Series C* **37**, 117-132.
- Fuller, W. A. (1987). *Measurement Error Models*. Wiley, New York.
- Gulliksen, H. (1950). *Theory of Mental Tests*. Wiley, New York.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663-685.
- Jowell, R. and the Central Co-ordinating Team. (2008). *European Social Survey 2008/2009: Technical Reports*. Centre for Comparative Social Surveys, City University, London.
- Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. Springer, New York.
- Lohr, S. L. (2010). *Sampling: Design and Analysis*, 2nd edition. Brooks/Cole, Boston.
- Lord, F. M. and Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley, Menlo Park, California.
- Mattei, F. and Niemi, R. G. (2005). Political efficacy. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), 525-534. Greenwood Press, Westport, Connecticut.
- Nathan, G. (1988). Inference based on data from complex sample designs. In *Handbook of Statistics*, Volume 6 (Sampling) (Edited by P. R. Krishnaiah and C. R. Rao), 247-266. North Holland, Elsevier, Amsterdam.
- Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society* **97**, 558-625.

-
- Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric Theory*, 3rd edition. McGraw-Hill, New York.
- Rudolph, T. J. (2005). Trust in government. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), 851-855. Greenwood Press, Westport, Connecticut.
- Skinner, C. J., Holt, D. and Smith, T. M. F. (1989). *Analysis of Complex Surveys*. Wiley, New York.
- StataCorp. (2001). *Stata Statistical Software: Release 7.0*. Stata Corporation, College Station, Texas.
- Thompson, M. E. (1997). *Theory of Sample Surveys*. Chapman and Hall, London.
- Valliant, R., Dorfman, A. H. and Royall, R. M. (1999). *Finite Population Sampling and Inference: A Prediction Approach*. Wiley, New York.

Received May 17, 2011; accepted November 8, 2011.

Gordon G. Bechtel
University of Florida and Florida Research Institute
P.O. Box 117155, Gainesville, Florida 32611-7155, USA
bechtel@ufl.edu