

True-Value Regression Theory

Gordon G. Bechtel

University of Florida and Florida Research Institute

Abstract: Design-based regression regards the survey response as a constant waiting to be observed. Bechtel (2007) replaced this constant with the sum of a fixed true value and a random measurement error. The present paper relaxes the assumption that the expected error is zero within a survey respondent. It also allows measurement errors in predictor variables as well as in the response variable. Reasonable assumptions about these errors over respondents, along with coefficient alpha in psychological test theory, enable the regression of true responses on true predictors. This resolves two major issues in survey regression, i.e. errors in variables and item non-response. The usefulness of this resolution is demonstrated with three large datasets collected by the European Social Survey in 2002, 2004 and 2006. The paper concludes with implications of true-value regression for survey theory and practice and for surveying large world populations.

Key words: Coefficient alpha, equation error, errors in variables, European Social Survey, Horvitz-Thompson estimation, imputed census, institutional trust, item non-response, variance estimation.

1. Placing Measurement Error in Design-Based Regression

Design-based sampling postulates the survey respondent to be in a constant state that may be observed on a rating scale such as 0 1 2 3 4 5 6 7 8 9 10. Thus, the value solicited on an opinion poll is regarded as a fixed number in waiting. More realistically, however, a survey rating should be regarded as an erroneous observation that departs from its true value. The present paper favors this more plausible view and generalizes Bechtel's (2007) interpretation of survey sampling.

Our unit of interest is an individual in a finite population. In a hypothetical census this individual's score on an attribute of interest is the sum of her (his) ratings on several items presumed to measure that attribute. A summative score is interpreted as a true attribute value plus a measurement error. This score error may contain fixed item biases, random item errors, and item imputation errors. Thus, the error in a respondent's score may not have zero expectation, and

her (his) expected score may not be her(his) true attribute value. In addition, measurement errors are allowed in predictor variables as well as the response variable. Reasonable assumptions about these errors *over respondents*, along with coefficient alpha in psychological test theory, enable the regression of true responses on true predictors.

The present treatment of errors in variables avoids model-based likelihood maximization that assumes 1) true predictor values to be random variables with a specified distribution and 2) apriori knowledge of measurement-error variance (Fuller, 1987; Whittemore and Keller, 1988). Here the distribution of true values over the population is unspecified, and error variance is *estimated* from information in the component ratings that make up a survey score. This provides a richer definition of measurement error. It also allows this error to be circumvented by Horvitz-Thompson type estimation of population totals.

The present paper also avoids model-based sampling, where a finite census of realized responses is itself a sample from a "superpopulation" with assumed distribution and covariance properties (Binder 1983; Nathan 1988; Skinner, Holt, and Smith 1989; Thompson 1997; Valliant, Dorfman, and Royall 1999). The present approach, in contrast, places measurement error into Neyman randomization theory. As noted by Bellhouse (1988, p.11),

A major motivating force behind many so-called classical sampling methods is the execution of large-scale social and economic surveys. There is a gradual move away from reliance on models for design and estimation toward a pure randomization approach so that the minimum of assumptions may be made about the population. As shown by Neyman (1934) sampling designs and estimation techniques based on incorrect models can lead to disastrous results.

Section 2 describes summative survey scores and extends the definition of measurement error. Section 3 distinguishes an imputed census from a finite population of true values and uses coefficient alpha in psychological test theory to specify measurement-error variance. This identifies parameters in the presence of both errors in variables and error in the regression equation. It also gives an estimable form for population regression coefficients. Section 4 estimates these coefficients using weighted sample totals that are estimates of corresponding population totals. Section 5 applies true-value theory to three large datasets collected by the European Social Survey in 2002, 2004 and 2006. Section 6 concludes the paper with implications of true-value regression for survey theory and practice and for surveying world populations.

2. Summative Attitude and Opinion Scores

2.1 The response score

Our response variable illustrating true-value regression is *institutional trust*, as measured by the European Social Survey (ESS). The ESS instructs the respondent as follows:

Using this card, please tell me on a score of 0-10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust. READ OUT... [country]'s parliament; the legal system; the police; politicians; the European Parliament; the United Nations¹. These six items were administered in 2002, 2004, and 2006. In our analysis individual *i*'s trust score is the sum of her (his) six responses and ranges from 0 to 60.

2.2 Predictor scores

We regress *institutional trust* on the five predictors in Table 1, which are core items in the ESS. The first two predictors are also multiple-item scores. The *institutional evaluation* score is the sum of five ratings and ranges from 0 to 50. The sum of the three ratings for *faith in people* ranges from 0 to 30. The other three predictors are fixed settings provided by each respondent's *interest in politics*, *age*, and *gender*.

Table 1: Predictors of the institutional trust score [0, 60]

Institutional evaluation score [0, 50]			
Economy in country	Extremely dissatisfied	0 1 2 3 4 5 6 7 8 9 10	Extremely satisfied
National government	Extremely dissatisfied	0 1 2 3 4 5 6 7 8 9 10	Extremely satisfied
How democracy works	Extremely dissatisfied	0 1 2 3 4 5 6 7 8 9 10	Extremely satisfied
Education in country	Extremely bad	0 1 2 3 4 5 6 7 8 9 10	Extremely good
Health services in country	Extremely bad	0 1 2 3 4 5 6 7 8 9 10	Extremely good
Faith in people score [0, 30]			
Most people	take advantage of you	0 1 2 3 4 5 6 7 8 9 10	try to be fair
With most people you	can't be too careful	0 1 2 3 4 5 6 7 8 9 10	can have trust
Most people	look out for themselves	0 1 2 3 4 5 6 7 8 9 10	try to be helpful
Interest in politics score [0,3]			
How interested would you say you are in politics?			
Not at all interested 0	Hardly interested 1	Quite interested 2	Very interested 3
Age [15, 102]			
Gender [0, 1]	Male 0	Female 1	

¹These instructions are found on the website ess.nsd.uib.no

2.3 True values and measurement errors in scores

Fixed and random measurement errors. For *institutional trust* respondent i 's observed rating on item m is

$$Y_{im} = \eta_{im} + E_{im} \quad \text{for } m = 1, \dots, 6,$$

where η_{im} is i 's true value on item m in the interval $[0, 10]$, and E_{im} is her (his) fixed bias or random error in responding to item m . Individual i 's derived trust score is

$$Y_i = \sum Y_{im} = \sum \eta_{im} + \sum E_{im} = \eta_i + E_i, \quad (2.1)$$

where the summations (over items) are over $m = 1, \dots, 6$. In (2.1) η_i is i 's true trust value in the interval $[0, 60]$, and E_i is the sum of i 's fixed and/or random item errors.

Referring to Table 1, let M_j be the number of items in predictor j 's summative scale score, where j denotes *institutional evaluation* or *faith in people*. In each case respondent i 's observed rating on item m is

$$X_{ijm} = \tau_{ijm} + U_{ijm} \quad \text{for } m = 1, \dots, M_j,$$

where τ_{ijm} is i 's true item value in the interval $[0, 10]$. The measurement error U_{ijm} is i 's fixed bias or random response error for item m in predictor j . Individual i 's derived score on the j -th predictor is

$$X_{ij} = \sum X_{ijm} = \sum \tau_{ijm} + \sum U_{ijm} = \tau_{ij} + U_{ij}, \quad (2.2)$$

where the summations (over items) are over $m = 1, \dots, M_j$. For *institutional evaluation* this score is in the interval $[0, 50]$. For *faith in people* it is in $[0, 30]$. In (2.2) τ_{ij} is i 's true predictor value, and U_{ij} is the sum of i 's fixed and/or random item errors.

The score errors $E_i = \sum E_{im}$ and $U_{ij} = \sum U_{ijm}$ in Equations (2.1) and (2.2) derive their characteristics from their item components E_{im} and U_{ijm} . This permits a more realistic diversity of measurement errors in scores. Thus, if E_i or U_{ij} contains only random item errors, each with zero expectation, then we have the classic situation in which E_i or U_{ij} has zero expectation. In this case the expectation of score Y_i or X_{ij} is its true value η_i or τ_{ij} . However, if E_i or U_{ij} a) is a mixture of fixed item biases and random item errors, b) contains only fixed item biases, or c) contains a random item error with non-zero expectation, then this score error has non-zero expectation for the i -th individual. Under any of these latter conditions the expectation of i 's score Y_i or X_{ij} is *not* her (his) true value η_i or τ_{ij} .

Imputation errors. The present theory reveals an imputation as a departure from a true value. Thus, imputation error is a special case of measurement error. If a Y_{im} in (2.1) or an X_{ijm} in (2.2) is an imputed item response, then E_{im} or U_{ijm} is an imputation error about the true item response η_{im} or τ_{ijm} . In this case a response or predictor score contains an imputation error, along with random errors and fixed biases associated with its other items.

3. Population Inference via an Imputed Census

Whether the E_i or U_{ij} in (2.1) or (2.2) contains random measurement errors, fixed item biases, item imputation errors, or mixtures of these for the i -th individual, this score error becomes a constant upon census realization. The values of the errorless predictors are also constants for the i -th individual. Letting $i = 1, \dots, N$, the realized census $\{Y_i, X_{i1}, \dots, X_{ip}\}$ of scores and errorless predictors is in one-to-one correspondence with a population $\{\eta_i, \tau_{i1}, \dots, \tau_{ip}\}$ of true attribute values and the same errorless predictors. (Note that $X_{ij} = \tau_{ij}$ for an errorless predictor j).

Our approach will be to make inferences to the population $\{\eta_i, \tau_{i1}, \dots, \tau_{ip}\}$ for $i = 1, \dots, N$. A window into this population is provided by the realized imputed census $\{Y_i, X_{i1}, \dots, X_{ip}\}$. This formulation requires imputed census item ratings that mimic actually imputed item ratings in our sample from this census. Section 5.3 describes our imputations of missing item responses in the ESS.

3.1 Parameter identification in the presence of equation error

Our population model is

$$\eta_i = \beta_0 + \beta_1 \tau_{i1} + \dots + \beta_p \tau_{ip} + \epsilon_i \quad \text{for } i = 1, \dots, N, \quad (3.1)$$

where ϵ_i is “error in the equation” (Fuller, 1987, p. 106) for fitting the true responses η_i by the true predictors $\tau_{i1}, \dots, \tau_{ip}$. In equation (3.1) the intercept β_0 , the slopes β_1, \dots, β_p , and the equation error ϵ_i are uniquely identified by the ordinary-least-squares (OLS) condition that $\sum \epsilon_i^2$ is minimal when the population of true responses η_i is regressed on the population of true predictors $\tau_{i1}, \dots, \tau_{ip}$. This OLS criterion, along with the inclusion of the intercept β_0 in (3.1), guarantees that the population sum $\sum \epsilon_i = 0$.

Our OLS identification of $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ is the function

$$\boldsymbol{\beta} = \left(\sum \mathbf{T}_i \mathbf{T}_i^T \right)^{-1} \sum \mathbf{T}_i \eta_i, \quad (3.2)$$

of true predictors and true responses. In (3.2) the two *population* summations (over individuals) run over $i = 1, \dots, N$, the row vector $\mathbf{T}_i^T = (\mathbf{1}, \tau_{i1}, \dots, \tau_{ip})$,

and N is the size of the population $\{\eta_i, \tau_{i1}, \dots, \tau_{ip}\}$. Equation (3.2) selects a unique parameterization $\beta_0, \beta_1, \dots, \beta_p, \epsilon_i$ from an infinite set of characterizations of the η_i .

In order to gain an estimable form of our target β it is necessary to specify the error variances of the erroneous predictors in the imputed census $\{Y_i, X_{i1}, \dots, X_{ip}\}$. This specification is given by coefficient alpha in psychological test theory (Cronbach, 1951; Lord and Novick, 1968; Nunnally and Bernstein, 1994; StataCorp., 2001).

3.2 Coefficient alpha as a census parameter

The census coefficient alpha for the j -th erroneous predictor is

$$\alpha_j = \frac{M_j}{M_j - 1} [1 - \sum \text{Var}(X_{jm}) / \text{Var}(X_j)], \quad (3.3)$$

where j denotes *institutional evaluation* or *faith in people*, and the summation (over items) runs over $m = 1, \dots, M_j$. In (3.3) X_{jm} and X_j are item and score variables whose variances are

$$\begin{aligned} \text{Var}(X_{jm}) &= [\sum X_{ijm}^2 - (\sum X_{ijm})^2 / N] / N \quad \text{and} \\ \text{Var}(X_j) &= [\sum X_{ij}^2 - (\sum X_{ij})^2 / N] / N, \end{aligned}$$

where the *census* summations (over individuals) run over $i = 1, \dots, N$. Lord and Novick (1968, pp. 87-89) showed that if measurement errors are uncorrelated with true scores and each other, then

$$\alpha_j \leq \text{Var}(\tau_j) / \text{Var}(X_j). \quad (3.4)$$

In the inequality (3.4) the variance of the true-value variable τ_j is

$$\text{Var}(\tau_j) = [\sum \tau_{ij}^2 - (\sum \tau_{ij})^2 / N] / N,$$

where the *population* summations (over individuals) run over $i = 1, \dots, N$. The right side of (3.4) is the ratio of the population variance $\text{Var}(\tau_j)$ of true scores to the census variance $\text{Var}(X_j)$ of observed and imputed scores for predictor j . If measurement errors U_{ij} in (2.2) are uncorrelated with true scores over the population and census, and the census sum $\sum U_{ij} = 0$, then

$$\text{Var}(X_j) = \text{Var}(\tau_j) + \text{Var}(U_j),$$

where $\text{Var}(U_j) = \sum U_{ij}^2 / N$ is the census variance of the error variable U_j . Finally, letting $\tau_{ijm} = \tau_{ij} / M_j$ in (2.2), we have

$$\begin{aligned} \alpha_j &= [\text{Var}(X_j) - \text{Var}(U_j)] / \text{Var}(X_j), \quad \text{or} \\ \text{Var}(U_j) / \text{Var}(X_j) &= 1 - \alpha_j, \end{aligned}$$

which gives

$$\sum U_{ij}^2 = (1 - \alpha_j)[\sum X_{ij}^2 - (\sum X_{ij})^2/N] = \delta_j. \tag{3.5}$$

In (3.5) the *census* summations (over individuals) run over $i = 1, \dots, N$. We now use δ_j to define estimable population regression coefficients.

3.3 An estimable form of the population target

In Section 3.2 we assumed that *over the population and census* realized measurement errors sum to zero, are uncorrelated with true scores, and are uncorrelated with each other. These assumptions may be written as

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

where j and k denote predictors. Under these assumptions, along with (2.1) and (2.2), it is easily shown that

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

Our population target (3.2) may then be written in the estimable form

$$\beta = (\sum X_i X_i^T - \Delta)^{-1} \sum X_i Y_i. \tag{3.6}$$

In (3.6) $X_i^T = (1, X_{i1}, \dots, X_{i5})$ and $\Delta = \text{diag}(0, \delta_1, \delta_2, 0, 0, 0)$, where δ_j for $j = 1, 2$ is given by (3.5). The first diagonal element δ_0 is zero because the intercept in (3.1) multiplies an error-free unit value. Also, $\delta_3 = \delta_4 = \delta_5 = 0$ because *interest in politics*, *age*, and *gender* in Table 1 are error-free variables.

4. Sampling and Estimation

The hypothetical census in Section 3 is an imputed census $\{Y_i, X_{i1}, \dots, X_{ip}\}$ in one-to-one correspondence with a complete population $\{\eta_i, \tau_{i1}, \dots, \tau_{ip}\}$ that has no missing values. Let a sample of n realizations be drawn without replacement from our census such that π_i for $i = 1, \dots, N$ is the probability of i 's inclusion in the sample. This sample provides estimates of the three components of β in (3.6) and hence an estimate of β in (3.2).

4.1 Estimating $\sum X_i X_i^T$ and $\sum X_i Y_i$

Complete data. Each element of the 6×6 matrix $\sum X_i X_i^T$ is the census sum of products $\sum X_{ij} X_{ik}$, and each element of the 6-vector $\sum X_i Y_i$ is the census sum of products $\sum X_{ij} Y_i$. In the absence of missing item responses in the census and sample, the Horvitz-Thompson (1952) type estimates of these census totals are given by $\sum w_i X_{ij} X_{ik}$ and $\sum w_i X_{ij} Y_i$, where the two *sample* summations (over individuals) run over $i = 1, \dots, n$. The weight w_i is the transform of $1/\pi_i$ described in Section 5.2. The matrix and vector estimates of $\sum X_i X_i^T$ and $\sum X_i Y_i$ are given by $\sum w_i X_i X_i^T$ and $\sum w_i X_i Y_i$ (Binder, 1983; Nathan, 1988, pp. 255-256; Thompson, 1997, pp. 106-107; Valliant, Dorfman, and Royall, 1999, pp. 40-41; Lohr, 1999, pp. 354-361; StataCorp., 2001, Volume 4, pp. 29-30; Chaudhuri and Stenger, 2005, pp. 264-265).

Missing data. In the presence of item non-response we assume imputed census item ratings that mimic actually imputed item ratings in our sample. These sample imputations replace the imputed census ratings that cannot be drawn in a sample. If an individual's sample score includes sample item imputations, it will differ slightly from her (his) census score that contains corresponding census imputations. In this case the weighted sample totals in the preceding paragraph will be slightly altered estimates of their census totals. This is because these weighted totals include individuals' sample scores that approximate their census scores. With large sample sizes these weighted totals differ negligibly from the totals that would be obtained with census imputations instead of sample imputations.

4.2 Estimating $\delta = \text{diag}(0, \delta_1, \delta_2, 0, 0, 0)$

Our estimate of predictor j 's coefficient alpha in (3.3) is

$$\mathbf{a}_j = \frac{M_j}{M_j - 1} \left[1 - \sum \text{Var}(X_{jm}) / \text{Var}(X_j) \right] \quad \text{for } j = 1, 2, \quad (4.1)$$

where the summation (over items) is over $m = 1, \dots, M_j$. In (4.1)

$$\begin{aligned} \text{Var}(X_{jm}) &= \left[\sum w_i X_{ijm}^2 - (\sum w_i X_{ijm})^2 / \sum w_i \right] / \sum w_i, \quad \text{and} \\ \text{Var}(X_j) &= \left[\sum w_i X_{ij}^2 - (\sum w_i X_{ij})^2 / \sum w_i \right] / \sum w_i, \end{aligned}$$

where the *sample* summations (over individuals) run over $i = 1, \dots, n$. The estimated alpha coefficients in Table 2 reveal that *institutional evaluation* is measured with higher reliability than *faith in people*. Substituting \mathbf{a}_j in (4.1) for α_j in (3.3)

gives our estimate of δ_j as

$$d_j = (1 - \mathbf{a}_j) \left[\sum w_i X_{ij}^2 - \left(\sum w_i X_{ij} \right)^2 / \sum w_i \right] \quad \text{for } j = 1, 2. \quad (4.2)$$

Table 2: Coefficient alphas for the erroneous predictors

	2002	2004	2006
Institutional evaluation	.784	.817	.803
Faith in people	.719	.719	.702

4.3 Estimating β

The estimator of β in (3.2) and (3.6) is

$$\mathbf{B} = \left(\sum w_i X_i X_i^T - D \right)^{-1} \sum w_i X_i Y_i, \quad (4.3)$$

where the two sample summations (over individuals) run over $i = 1, \dots, n$. In (4.3) $D = \text{diag}(0, d_1, d_2, 0, 0, 0)$ where d_j for $j = 1, 2$ is given by (4.2). The first diagonal element d_0 of D is zero because the intercept in (3.1) multiplies an errorless unit value. Also, $d_3 = d_4 = d_5 = 0$ because the last three predictors in Table 1 are without error. Note that under the inverse in (4.3) d_1 and d_2 correct the well-known design-based regression formula which holds when $D = 0$.

The estimated standard errors of the coefficients B_0, B_1, \dots, B_p of \mathbf{B} in (4.3) are derived in the appendix. These estimates contain the effects of errors in variables on coefficient variance. They appear in Table 3 for our regressions from the first three rounds of the ESS.

5. The European Social Survey of Institutional Trust

5.1 Repeated surveys

Repeated national surveys in the United States are described by Firebaugh (2005, pp. 680-681): Repeated surveys pose the same questions to different samples of people . . . repeated surveys track aggregate change over time- for example, change in the percentage of the American electorate who identify themselves as Democrats. Repeated surveys have become a staple of social science research. Perhaps the best known repeated survey in the United States is the General Social Survey (GSS) . . . To monitor attitudinal change in the United States, the GSS selects a fresh sample for each new survey. The present paper regresses scores derived from repeated *cross-national surveys*. The data for our analysis was supplied by the European Social Survey (Jowell and the Central

Co-ordinating Team, 2003; 2005; 2007) which is among the first social science projects to receive funding to support its infrastructure and in 2005 was awarded Europe's top annual science award, the Descartes prize.

The ESS is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. Now moving into its fourth round, the survey covers over 30 nations and employs the most rigorous methodologies. The survey has been 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 variables².

5.2 Sampling and euoweighting

The present study uses items from the ESS core module that were administered in 2002, 2004, and 2006. In order to monitor *institutional trust*, our analysis is limited to twelve EU nations that were surveyed in all three rounds of the ESS. These EU members are Great Britain, France, Germany, Spain, Portugal, Belgium, Denmark, Sweden, Finland, Poland, Hungary, and Slovenia.

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. $n_{net} = n_{eff} * DEFF = 1,500 * DEFF$. You will get an estimate for the design effect DEFF in your country from your sampling expert.
- The gross sample size is calculated as follows: $n_{gross} = n_{net} / (RR * ER)$, where RR (target is 70%) is the predicted response rate and ER is the eligibility rate³.

ESS provides design weights and population size weights to construct weights representative of the European population. Design weights are normed inverses

²These passages are quoted from the website www.europeansocialsurvey.org.

³These three specifications appear on the website www.europeansocialsurvey.org.

of sample inclusion probabilities, and they sum to each country's net sample size. A country's population size weight is

$$(\text{population size aged 15 and over})/(\text{net sample size in the data file} \times 10000).$$

Then

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

insures that each of our three weighted regressions represents a country in proportion to its population size. An individual's euroweight is a transformed inverse probability ($1/\pi_i$) of her (his) inclusion in the ESS sample.

ESS notes that its design weights, and therefore its national- and euro-weighting procedures, do not adjust for *unit non-response* in its national samples⁴. In the present analysis *item non-response* is treated by imputations described in Section 5.3. The theory supporting these imputations is given in Sections 2 through 4.

5.3 Within-scale imputation

Sample loss for a multiple-item score is avoided here by regression imputations confined to those items making up that score. For example, five imputations were carried out for the 5-item institutional evaluation scale in Table 1. Each item in this predictor was regressed on the four other items, as well as age and gender, to impute its missing responses (cf. StataCorp., 2001, Volume 2: 69-71). Such within-scale imputations were computed for each of the two multiple-item scales in Table 1, as well as for the 6-item response scale described in Section 2.1. The error-free variables *interest in politics*, *age* and *gender* were not imputed. This procedure preserved 99% of the 2002, 2004, and 2006 samples. The number of cases for each of our three regressions is given at the bottom of Table 3.

5.4 Results for true-value regressions

Monitoring trust level. True value theory enables inferences to population means as well as regression coefficients. Summing both sides of (2.1) over $i = 1, \dots, N$ and using $\sum E_i = 0$ gives

$$\sum Y_i/N = \sum \eta_i/N. \quad (5.1)$$

The left side of (5.1) is an estimable form of the true population mean on the right side. Our estimator of this true mean is

$$m = \sum w_i Y_i / \sum w_i, \quad (5.2)$$

⁴This note may be found on the website ess.nsd.uib.no

where the two *sample* summations (over individuals) run over $i = 1, \dots, n$.

The estimated means in Table 3 were calculated with the euro weights in Section 5.2 replacing the w_i in equation (5.2). These estimates show that Europeans' trust in their institutions dropped from 2002 to 2004 but rebounded by 2006.

Monitoring effects on trust. For each ESS survey the trust score Y_i in (2.1) was regressed on the five predictors in Table 1 over our 12-nation aggregate. The regression coefficients, below the third line in Table 3, were computed from formula (4.3). Table 3 shows that women and younger respondents express greater trust in European institutions. The effects of *gender*, *age* and *institutional evaluation* remained stable over the first three rounds of the ESS. However, *interest in politics* declined between 2002 and 2004, and *faith in people* weakened between 2004 and 2006. The R-squares in Table 3 indicate that almost half of the variation in *institutional trust* is attributable to the five predictors in Table 1.

Table 3: Means and regression coefficients for institutional trust

	2002	2004	2006
Mean	28.68 (.099)**	27.65 (.100)**	28.01 (.091)
Number of cases	22428	22177	23235
Institutional evaluation	.810 (.019)	.813 (.018)	.849 (.018)
Faith in people	.380 (.035)	.436 (.037)**	.321 (.038)
Interest in politics	2.124 (.099)**	1.600 (.095)	1.631 (.086)
Age	-.028 (.004)	-.017 (.004)	-.017 (.004)
Gender	.938 (.163)	.927 (.155)	1.324 (.149)
R-squared	.442	.494	.480
Number of cases	22239	22009	23134

NOTE: The means and regression coefficients are weighted by the euro weights described in Section 5.2. Their standard errors are in parentheses. All coefficients below the third line are significant beyond the .00004 level. A double star (**) indicates a change between successive surveys that is significant beyond the .03 level. All other changes are not significant.

5.5 Comparisons to naive observed-value regressions

Equations (5.1) and (5.2) show that the means in Table 3 estimate both the true population mean and the realized census mean of *institutional trust*. However, the naive regression coefficients in Table 4 are biased by errors in the *institutional-evaluation* and *faith-in-people* scores. The downward bias in coefficients of erroneous predictors is well known (Johnston, 1984, pp. 428-430). However, the direction of bias, if any, in the slopes of error-free predictors cannot

be anticipated. Table 4 reveals that the *gender* effect is biased downward by the two erroneous predictors and that the *age* and *interest- in-politics* effects are unaffected.

Table 4 also shows that errors in variables spuriously reduce the standard errors of naive regression coefficients. This contributes to the misleading “change” in the effect of *institutional evaluation* between 2002 and 2004. Table 3 shows that this effect remained stable over the first three rounds of the ESS.

Table 4: Naive regression coefficients for institutional trust

	2002	2004	2006
Institutional evaluation	.637 (.010)**	.671 (.010)	.679 (.009)
Faith in people	.362 (.017)	.399 (.016)**	.328 (.016)
Interest in politics	2.139 (.094)**	1.677 (.090)	1.674 (.082)
Age	-.027 (.004)	-.017 (.004)	-.018 (.004)
Gender	.804 (.159)	.839 (.152)	1.159 (.144)

NOTE: All coefficients are significant beyond the .00004 level. A double star (**) indicates a change between successive surveys that is significant beyond the .02 level. All other changes are not significant.

6. Implications of True-Value Regression

6.1 Survey theory and practice

True-value theory departs from psychological test theory (Lord and Novick, 1968) by allowing the error in a survey respondent’s score to have nonzero expectation. This score error may contain fixed item biases, random item errors, and item imputation errors. Therefore, a respondent’s expected score on a survey scale is no longer restricted to her (his) true attribute value.

True-value theory also relaxes the strong assumption in design-based sampling theory that a finite population is a set of errorless constants (Nathan, 1988; Lehmann, 1999; Lohr, 1999; Chaudhuri and Stenger, 2005). Here this classical population is replaced by *two* finite sets of real vectors. The first set is a population of N true vectors, and the second set is a census of N realized vectors. The interpretation of realizations Y_i and X_{ij} in (2.1) and (2.2) as deviations from true values η_i and τ_{ij} , rather than errorless constants, is a step forward in the Neyman paradigm (cf. Bellhouse, 1988). This step extends randomization theory without invoking an abstract superpopulation postulated by model-based theory (cf. Skinner, Holt, and Smith, 1989). Thus, true-value theory establishes a middle ground between design- and model-based regression.

Finally, true-value theory resolves two major issues in survey regression by 1) allowing errors in predictor variables and 2) viewing imputation for item non-response as a special case of measurement error. Reasonable assumptions about the behavior of measurement error *over the population and census*, along with coefficient alpha in psychological test theory, enable the regression of true responses on true predictor values. Tables 3 and 4 show that our estimator in (4.3), which is corrected for errors in predictors, makes a considerable difference in monitoring large cross-national populations.

6.2 Surveying institutional trust in the 21st century

True-value regression is demonstrated here with a response variable, European *institutional trust*, defined in Section 2.1. The importance of this survey variable is seen in the following review of Leonard's (2005) hypothesis:

What Europe has, argues Mark Leonard in his provocatively titled book, *Why Europe Will Run the 21st Century*, is a model, one centered around a new understanding of power and embodied in the institutions and norms of the European Union. The EU exerts an irresistible attraction on the countries around it, Leonard says, drawing them into its orbit, embedding them in its legal and economic framework and changing them from the inside out. Next to this "transformative power," the United States' military might, which can change regimes but not societies, and whose application is necessarily fleeting, seems a weak instrument indeed. Increasingly, Leonard tells us, we'll see more regional groupings emerge bound, as the EU is, by mutual self-interest and common values. It's in this sense, he argues, that Europe-or, more precisely, the "European way"-will dominate the 21st century⁵.

Whether or not the "European way" leads 21st century public opinion, it is clear that the accurate monitoring of institutional trust in the EU, China, and the United States is crucial to each of these nations and the rest of the world. Intense distrust of the Iraq war in America and Europe has been evident for several years. More recently, there has been worldwide distrust with institutional solutions to the global financial crisis. American opinion surveys suggest that: Feelings of trust in government are grounded in one's assessment of how well government is satisfying normative expectations for government performance . . . Numerous studies report that citizens' level of trust in government is linked to various indicators of government performance (Rudolph, 2005, pp. 851-853). An indicator of government performance in American surveys is presidential approval. This is measured by The Gallup Organization, Zogby International, CBS News/New York Times, ABC/Washington Post, NBC News/Wall Street Journal, and the

⁵This quotation has been taken from the website www.motherjones.com

American National Election Studies (Clarke, Stewart, and Rodgers, 2005). In an analysis of Zogby data Bechtel (2007) reported that jobs, the economy, and the Iraq war were the strongest predictors of presidential approval. A second indicator of government performance in the United States is congressional approval.

Congressional approval is the extent to which the American public approves of the job that congress, as a whole, is doing at any given point in time. These public evaluations are generally based both on what Congress is doing, in terms of representing the public and keeping the country running smoothly, and on how it is doing it. Congressional approval is an important concept in political science as it is believed to influence the public's trust in our system of government in general (McDermott, 2005, p.115). Approval of European institutional performance is measured in the ESS by the *institutional evaluation* score in Table 1. This score taps satisfaction with the performance of economic, educational, and health service sectors, as well as satisfaction with a country's national government and how it works. In line with the American link between government performance and trust in government, Table 3 shows that public approval of EU institutions is a consistent predictor of trust in those institutions.

The European Social Survey, launched in 2002 with firm multi-national infrastructures, carries promise for understanding and monitoring institutional trust in the 21st century. The ESS has extended the objects of trust to include the European Parliament and the United Nations, along with the four national institutions listed in Section 2.1. The resulting 6-item trust score is computed from equation (2.1). The measurement errors in (2.1) and (2.2), as deviations from true values, correct extreme interpretations of this trust score and its predictors as errorless constants or specifically distributed random variables. Tables 3 and 4 show that this more realistic interpretation of micro data reduces bias in attitude and opinion regressions that make inferences to large world populations.

Acknowledgements

This work is dedicated to the direction given survey research by Daniel Horvitz and Donovan Thompson. It is a generalization of Bechtel (2007) and was supported by the University of Florida's Warrington College of Business Administration. The 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 in the present paper may be attributed to the European Social Survey or the University of Florida. The author thanks the JDS editor and reviewer for improving the presentation of true-value theory and its application.

Appendix. Variance Estimation for True-Value Regression

A.1 Population, census, and sample estimating functions

In the regressions in Table 3 the true values of *institutional trust*, *institutional evaluation* and *faith in people* are not observable. Nevertheless, we state the population estimating function for β in (3.2) in terms of these true values as

$$G(\beta) = \sum \mathbf{T}_i \eta_i - \left(\sum \mathbf{T}_i \mathbf{T}_i^T \right) \beta,$$

where the two *population* summations (over individuals) run over $i = 1, \dots, N$. Under the assumptions that measurement errors are uncorrelated with true values and each other *over the population and census*, $G(\beta)$ can be written as the census estimating function

$$G(\beta) = \sum X_i Y_i - \left(\sum X_i X_i^T - \Delta \right) \beta = \sum X_i (Y_i - X_i^T \beta) + \Delta \beta,$$

where Δ is given in (3.6), and all *census* summations (over individuals) run over $i = 1, \dots, N$. For a sample from this census, the sample estimating function for β is

$$g(\beta) = \sum w_i X_i Y_i - \left(\sum w_i X_i X_i^T - D \right) \beta = \sum w_i X_i (Y_i - X_i^T \beta) + D \beta, \quad (\text{A.1})$$

where the *sample* summations (over individuals) in (A.1) run over $i = 1, \dots, n$.

A.2 The variance of \mathbf{B} in (4.3)

A first-order Taylor expansion of $g(\mathbf{B})$ at $\mathbf{B} = \beta$, where β is the population target in (3.6), gives the approximation

$$g(\mathbf{B}) \approx g(\beta) + \frac{\partial g(\beta)}{\partial \beta} (\mathbf{B} - \beta)$$

(cf. Binder, 1983; Thompson, 1997, pp. 106-111; Lohr, 1999, pp. 290-293; StataCorp., 2001, Volume 4, pp. 29-31). Because $g(\mathbf{B}) = 0$, we have

$$g(\beta) \approx - \frac{\partial g(\beta)}{\partial \beta} (\mathbf{B} - \beta),$$

which linearizes our estimation error as

$$\mathbf{B} - \beta \approx - \left(\frac{\partial g(\beta)}{\partial \beta} \right)^{-1} g(\beta).$$

Therefore,

$$\begin{aligned} \text{Var}(\mathbf{B}) &\approx \left(\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^{-1} \text{Var}[g(\boldsymbol{\beta})] \left(\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^{-1} \\ &= A^{-1} \text{Var}[g(\boldsymbol{\beta})] A^{-1}, \end{aligned} \quad (\text{A.2})$$

where $A = \sum w_i X_i X_i^T - D$.

A.3 Iterative estimation of $\text{Var}(\mathbf{B})$

We partition the variance of $g(\boldsymbol{\beta})$ in (A.1) as

$$\begin{aligned} \text{Var}[g(\boldsymbol{\beta})] &= \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \boldsymbol{\beta}) + D\boldsymbol{\beta}\right] \\ &= \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \boldsymbol{\beta})\right] + \text{Var}[D\boldsymbol{\beta}]. \end{aligned}$$

Replacing $\boldsymbol{\beta}$ by \mathbf{B} in each variance component gives our estimate of $\text{Var}[g(\boldsymbol{\beta})]$ as

$$\text{Var}[g(\boldsymbol{\beta})] = \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \mathbf{B})\right] + \text{Var}[D\mathbf{B}]. \quad (\text{A.3})$$

The first matrix on the right of (A.3) has the same form as the well-known covariance matrix in design-based regression (Lohr, 1999, p. 360; Bechtel, 2007). Its computation is described in StataCorp. (2001, Volume 4, pp. 29-30, 69-71). Using (A.3) our estimate of $\text{Var}(\mathbf{B})$ in (A.2) is

$$\begin{aligned} \text{Var}(\mathbf{B}) &= A^{-1} \text{Var}[g(\boldsymbol{\beta})] A^{-1} \\ &= A^{-1} \left\{ \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \mathbf{B})\right] + \text{Var}[D\mathbf{B}] \right\} A^{-1} \\ &= A^{-1} \left\{ \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \mathbf{B})\right] \right\} A^{-1} + A^{-1} D \text{Var}(\mathbf{B}) D A^{-1}, \end{aligned}$$

which sets up our algorithm for solving for $\text{Var}(\mathbf{B})$. This algorithm is initialized by setting

$$\text{Var}_0(\mathbf{B}) = A^{-1} \left\{ \text{Var}\left[\sum w_i X_i (Y_i - X_i^T \mathbf{B})\right] \right\} A^{-1} = \mathbf{K},$$

which gives

$$\text{Var}_1(\mathbf{B}) = \mathbf{K} + A^{-1} D \mathbf{K} D A^{-1}.$$

The next iteration is computed as

$$\text{Var}_2(\mathbf{B}) = \mathbf{K} + A^{-1} D \text{Var}_1(\mathbf{B}) D A^{-1},$$

and the iterations are continued until

$$\text{Var}_q(\mathbf{B}) \approx \mathbf{K} + A^{-1} D \text{Var}_{q-1}(\mathbf{B}) D A^{-1}. \quad (\text{A.4})$$

Upon convergence, which occurs for $q \leq 10$, we estimate $Var(\mathbf{B})$ by $Var_q(\mathbf{B})$ in (A.4). The square roots of the $(p + 1)$ diagonals of $Var_q(\mathbf{B})$ are the estimated standard errors of the coefficients B_0, B_1, \dots, B_p of \mathbf{B} in (4.3).

References

- Bechtel, G. G. (2007). Distribution-free regression: Reinterpreting design-based sampling. *Journal of Data Science* **5**, 535-554.
- 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) pp. 1-14, North Holland.
- Binder, D. A. (1983). On the variances of asymptotically normal estimators from complex samples. *International Statistical Review* **51**, 279-292.
- Chaudhuri, A. and Stenger, H. (2005). *Survey Sampling: Theory and Methods*. 2nd edition. Chapman and Hall.
- Clarke, H. D. , Stewart, M. C. and Rodgers, C. (2005). Presidential approval. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), pp. 571-579, Greenwood Press.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* **16**, 297-334.
- Firebaugh, G. (2005). Repeated surveys. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), pp. 680-683, Greenwood Press.
- Fuller, W. A. (1987). *Measurement Error Models*. Wiley.
- 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.
- Johnston, J. (1984). *Econometric Methods*, 3rd edition. McGraw-Hill.
- Jowell, R. and the Central Co-ordinating Team. (2003, 2005, 2007). European Social Survey 2002/2003, 2004/2005, 2006/2007: Technical Reports. Centre for Comparative Social Surveys, City University, London.
- Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. Springer.
- Leonard, M. (2005). *Why Europe Will Run the 21st Century*. Harper-Collins.
- Lohr, S. L. (1999). *Sampling: Design and Analysis*. Duxbury Press.
- Lord, F. M. and Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley.

- McDermott, M. L. (2005). Congressional approval. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), pp. 115-118, Greenwood Press.
- 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), pp. 247-266, North Holland.
- Neyman, J. (1934). On 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.
- Rudolph, T. J. (2005). Trust in government. In *Polling America: An Encyclopedia of Public Opinion* (Edited by S. J. Best and B. Radcliff), pp. 851-855, Greenwood Press.
- Skinner, C. J. , Holt, D. and Smith, T. M. F. (1989). *Analysis of Complex Surveys*. Wiley.
- StataCorp. (2001). *Stata Statistical Software: Release 7.0*. Stata Corporation.
- Thompson, M. E. (1997). *Theory of Sample Surveys*. Chapman and Hall.
- Valliant, R. , Dorfman, A. H. and Royall, R. M. (1999). *Finite Population Sampling and Inference: A Prediction Approach*. Wiley.
- Whittemore, A. S. and Keller, J. B. (1988). Approximations for regression with covariate measurement error. *Journal of the American Statistical Association* **83**, 1057-1066.

Received March 7, 2009; accepted June 14, 2009.

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