Who Changed their Minds about Erap? Hypothesis Tests based on Design-Consistent Estimates of an Ordered Probit Model of the Philippine President's Ratings

Michael M. Alba, Geriz Bigol, Reina Filipina Deacosta, and Ria Pulumbarit*

Abstract

Using data from the March and June 1999 surveys of the Social Weather Stations (SWS), we generate and analyze design-based estimates of the subpopulation means and standard deviations of President Estrada's ratings and of his net satisfaction index over certain respondent characteristics. We also estimate the parameters of a design-consistent ordered probit model of the ratings and compare the coefficient estimates as well as the marginal effects on the ratings probabilities of the March and June regressors to identify the respondent attributes that contributed to the net decline in his popularity. We find that although the estimates of the population means of the president's ratings and of the net satisfaction index were not statistically different between the two survey periods, by June 1999, residents of the Visayas as well as persons with some education had become disenchanted with the Estrada administration. Only for residents of Luzon outside the National Capital Region did the satisfaction ratings turn in favor of President Estrada.

I. Introduction

Filipinos are fascinated with politics. Indeed, politics is said to be the national pastime of the country. Hence, it is all the more surprising that, in the Philippines, the application of the science of politics—particularly in the employment of methods of statistical estimation and inference—lags behind the practice of the art. A case in point is the brouhaha that met the drop in President Joseph Estrada's ratings in the opinion polls in 1999. As reported in the Social Weather Stations' (SWS) website, the president's net satisfaction index took a tailspin from its peak of 67 percentage points in March 1999 to 65 percentage points in June, 28 percentage points in October, and 5 percentage points in December of the same year.[1]

For all the noise and cacophony of voices that the poll results precipitated, however, the discussions and debates in the media were marked nonetheless by a disconcerting dearth of analysis on one important issue: the demographic and socioeconomic characteristics of those individuals whose satisfaction with the president's performance had turned southward. More precisely, no systematic analysis seems to have been undertaken on who had changed their minds about Erap between surveys—when presumably the information would have been

* Respectively, associate professor of economics and graduates of the AB Economics program, De La Salle University, 2401 Taft Avenue, Manila 1004. We wish to thank the Social Weather Stations (SWS) and Mahar Mangahas and Jeanette Ureta, in particular, for giving us access to the data sets of the March and June 1999 surveys. We are also grateful to Edward Korn and Barry Graubard for kindly responding to our queries on design-consistent variance estimates and test statistics, Bryan Sayer for pointing us to the work of Korn and Graubard, and participants of seminars at the University of the Philippines School of Economics and the SWS.

[1] The net satisfaction index is simply the difference between the proportion of Filipinos of voting age who are satisfied (including those who are very satisfied) with the president's performance in office and the proportion of those who are dissatisfied (including those who are very dissatisfied), expressed in percentage points.

invaluable in formulating initiatives that would arrest the ratings decline or would recapture the hearts and minds of Filipinos who had become disgruntled with the Estrada administration.

For the subject to have been adequately studied, at least two concerns of statistical inference needed to be addressed. The first and more basic nicety stems from the fact that the data collection procedures of opinion polls tend to have complex survey designs (that involve stratification, multistage clustering, and differential selection probabilities of respondents), in large part to drive down survey costs as well as to reduce the logistical effort. A consequence of this strategy, however, is that design-based statistical estimators have to be used in lieu of the usual estimators that are based on the simple random sampling method. Otherwise, considerably smaller estimates of the standard errors are apt to be reported, which in turn would lead to the wrong inferences being made.

The second and more subtle issue involves the utilization of statistical two-way frequency tables *as the sole bases for drawing inferences*. As is well-known in classical econometrics (see Greene (2000), for example), a problem with such parsimonious models (in which the analysis focuses only on two or three variables at a time) is that they are apt to be infected with omitted variable bias.[2] Hence, a better methodological strategy is to generate such two-way tables or summary statistics as part of the preliminary exploratory data analysis, but to also estimate parametric models that allow the independent effect of an explanatory variable to be isolated from those of other explanatory variables.

In this paper, we take a second, more careful look at the decline in President Estrada's satisfaction ratings in 1999. Using data from the March and June 1999 surveys of the SWS,[3] we present and analyze some statistics of his ratings among respondents and certain of their characteristics, in which the estimators used are consistent with the multistage sampling design of the surveys. In addition, we separately estimate a design-based ordered probit model of the president's ratings in each survey period and develop design-consistent hypotheses tests that allow us to compare (a) the parameter estimates of the ordered probit model and (b) the marginal effects of the regressors on the probabilities of the ratings. In turn, these tests enable us to identify those respondent characteristics that contributed to the net change in the president's net satisfaction index. We find that although the estimates of the population means of the president's ratings and of the net satisfaction index were not statistically different between the two survey periods, by June 1999, residents of the Visayas as well as persons with some education had become disenchanted with the Estrada administration. Only for residents of Luzon outside the National Capital Region did the satisfaction ratings turn in favor of President Estrada.

The rest of the paper is organized as follows: Section II describes the empirical framework from which the ordered probit model is derived and presents the design-consistent statistical tests used to determine whether or not the parameter estimates and the marginal effects of the regressors on the ratings probabilities are significantly different between the two survey periods. This is followed in Section III by a brief discussion of the data sets and their sampling

---

[2] Omitted variable bias results when an explanatory variable in the true model is omitted in the estimated model and it happens to be correlated with some of the included explanatory variables. The omission has the effect of showing spurious correlations between the included variables and the dependent variable.

[3] As a matter of policy, the SWS imposes an embargo on data sets that are less than a year old. At the time the research for this paper was undertaken, the March and June 1999 data sets were the latest ones that were made available to researchers.

design parameters as well as by an introduction of the variables used. In Section IV, we present and interpret the statistical tables and the parameter estimates of our ordered probit model as well as the results of the tests of hypotheses. Section V concludes the paper by summarizing its findings.

II. Empirical Model

Consider a Filipino of voting age (who is chosen as a respondent in an opinion survey). Presumably, as a rational individual, he has an index function, U^*, which is an ordinal measure of his satisfaction level. In particular, let the value of his index function indicate his disposition to the state of affairs in the country (including the president's performance in office), such that the higher the value of U^*, the more favorably disposed he is.

We assume that U^* is fully known to the person himself, but not to others (particularly nosey statisticians and researchers). Suppose, however, that U^* is a linear combination of the respondent's observed and unobserved characteristics, so that it may be written as

$$U^* = \boldsymbol{\beta}'\mathbf{x} + \varepsilon, \tag{1}$$

where $\mathbf{x}$ is the vector of the observed factors, ε reflects the net influence of the unobserved factors on U^*, and $\boldsymbol{\beta}$ is the vector of coefficients. The index variable U^* remains a latent variable for researchers and statisticians, since neither $\boldsymbol{\beta}$ nor ε is observed by them.

In lieu of U^*, what may be gleaned from an opinion survey is the value of a polychotomous categorical variable, U, on the respondent's alleged disposition, where

$$U = \begin{cases} 1 & \text{if the respondent is very dissatisfied,} \\ 2 & \text{if he is somewhat dissatisfied,} \\ 3 & \text{if he is neither dissatisfied nor satisfied,} \\ 4 & \text{if he is somewhat satisfied,} \\ 5 & \text{if he is very satisfied.} \end{cases} \tag{2}$$

Nonetheless, U^*, being an ordinal index, can be appropriately scaled so that a one-to-one correspondence would exist between U^* and U.[4] For instance, it may be specified that

$$U = \begin{cases} 1 & \text{if } \quad U^* \le c_1, \\ 2 & \text{if } c_1 < U^* \le c_2, \\ 3 & \text{if } c_2 < U^* \le c_3, \\ 4 & \text{if } c_3 < U^* \le c_4, \\ 5 & \text{if } c_4 < U^*, \end{cases} \tag{3}$$

[4]More precisely, U^* may be taken as a positive monotonic transformation of the person's original ordinal index function that is appropriately scaled to have a one-to-one correspondence with U. As a monotone transform, U^* will continue to preserve the ranking or ordering of the person's preferences.

where c_i, for $i = 1, 2, 3, 4$, are arbitrary cut-off points in the person's *preference space*, such that $c_1 < c_2 < c_3 < c_4$.

Assuming the disturbance term, ε, to be independent of **x** and to be normally and independently distributed across the survey respondents (with zero mean and unit variance), the researcher can begin to make probabilistic statements that relate a person's observed characteristics to his rating of the national state of affairs. For example, the researcher can posit that

$$\begin{aligned} P(U=1) &= P(U^* \le c_1) \\ &= P(\boldsymbol{\beta}'\mathbf{x} + \varepsilon \le c_1) \\ &= P(\varepsilon \le c_1 - \boldsymbol{\beta}'\mathbf{x}) \\ &= \Phi(c_1 - \boldsymbol{\beta}'\mathbf{x}), \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function. Hence, in general, the probability that a particular state of disposition i is chosen may be derived as

$$P(U=i) = \begin{cases} \Phi(c_1 - \boldsymbol{\beta}'\mathbf{x}) & \text{if } i=1, \\ \Phi(c_2 - \boldsymbol{\beta}'\mathbf{x}) - \Phi(c_1 - \boldsymbol{\beta}'\mathbf{x}) & \text{if } i=2, \\ \Phi(c_3 - \boldsymbol{\beta}'\mathbf{x}) - \Phi(c_2 - \boldsymbol{\beta}'\mathbf{x}) & \text{if } i=3, \\ \Phi(c_4 - \boldsymbol{\beta}'\mathbf{x}) - \Phi(c_3 - \boldsymbol{\beta}'\mathbf{x}) & \text{if } i=4, \\ 1 - \Phi(c_4 - \boldsymbol{\beta}'\mathbf{x}) & \text{if } i=5. \end{cases} \tag{5}$$

Given a cross-section data set of sample size n generated by simple random sampling, the researcher may estimate the parameters, $\boldsymbol{\beta}$, c_1, c_2, c_3, and c_4, of this ordered probit model by the method of maximum likelihood using the following log-likelihood function:

$$\ln L = \sum_{t=1}^{n} \sum_{i=1}^{5} k_{ti} \ln P(U_t = i),$$

where

$$k_{ti} = \begin{cases} 1 & \text{if } U_t = i, \quad \text{for } i = 1, \ldots, 5, \text{and } t = 1, \ldots, N, \\ 0 & \text{otherwise.} \end{cases}$$

Since our data sets are generated by multistage sampling procedures with stratification clustering, and differential selection probabilities, however, the parameter estimates have to be based on the following *pseudo-log-likelihood* function instead:[5]

$$\ln L^* = \sum_{s=1}^{S} \sum_{p=1}^{P_s} \sum_{t=1}^{T_{sp}} w_{spt} \sum_{i=1}^{5} k_{spti} \ln P(U_{spt} = i),$$

where $s = 1, \ldots, S$ are the strata, $p = 1, \ldots, P_s$ are the primary sampling units (PSUs) in stratum s, $t = 1, \ldots, T_{sp}$ are the respondents in PSU p of stratum s, w_{spt} is the sample weight of the tth respondent in PSU p of stratum s, and

[5] It is so-called because it is derived from a weighted likelihood function, which is not the distribution function of the sample and therefore is not a true likelihood function.

$$k_{spti} = \begin{cases} 1 & \text{if } U_{spt} = i, \quad \text{for } i = 1,\ldots,5, \\ 0 & \text{otherwise.} \end{cases}$$

The object of this paper, though, is not to estimate the parameters of the ordered probit model per se, but to identify the regressors that contributed to the decline in the president's ratings between March and June 1999. In line with this goal, we implement two sets of hypothesis tests. The first compares the parameter estimates of the separately estimated ordered probit models for the same regressor (to identify the factors that underwent significant changes in their index function weights, $\boldsymbol{\beta}$), and the second compares the marginal effects of each regressor on the probabilities of the president's ratings (to identify the factors that underwent significant changes in their effects on the probabilities).

We use a Wald test statistic to evaluate the null hypothesis that $(\beta_{j,\text{March}} = \beta_{j,\text{June}})$ for each regressor j. Since the covariances of the March and June parameter estimates are zero (given that their data sets are independently generated), the Wald test statistic may be specified as

$$T_1 = \frac{(\hat{\beta}_{j,\text{March}} - \hat{\beta}_{j,\text{June}})^2}{\hat{V}(\hat{\beta}_{j,\text{March}}) + \hat{V}(\hat{\beta}_{j,\text{June}})} \sim F_{1,d_s}, \tag{6}$$

where

$$d_s = \frac{[\hat{V}(\hat{\beta}_{j,\text{March}}) + \hat{V}(\hat{\beta}_{j,\text{June}})]^2}{\dfrac{\hat{V}(\hat{\beta}_{j,\text{March}})^2}{d_{\text{March}}} + \dfrac{\hat{V}(\hat{\beta}_{j,\text{June}})^2}{d_{\text{June}}}}$$

is the degrees of freedom associated with $V(\hat{\beta}_{j,\text{March}} - \hat{\beta}_{j,\text{June}})$ based on a Satterthwaite approximation with d_q being the difference between the number of sampled PSUs in survey period q and the number of strata in survey period q, for q = March, June.

On the other hand, the derivation of the test statistic for evaluating the difference in the marginal effects of each regressor on the probability that $U = i$ is a bit more involved for two reasons: First, our marginal effects are atypical in that our regressors are sets of mutually exclusive dummy variables, such as area of residence, sex, age group, highest educational level completed, marital status, socioeconomic class, and religious affiliation. Thus, in the case of the specification at hand, the marginal effects on the probabilities of setting the jth independent variable equal to one may be given as follows:

$$\left.\frac{\Delta\hat{P}(U=i)}{\Delta x_j}\right|_{\mathbf{x}=\mathbf{0}} = \begin{cases} \Phi(\hat{c}_1 - \hat{\beta}_j) - \Phi(\hat{c}_1) & \text{if } i = 1, \\ \Phi(\hat{c}_2 - \hat{\beta}_j) - \Phi(\hat{c}_2) - [\Phi(\hat{c}_1 - \hat{\beta}_j) - \Phi(\hat{c}_1)] & \text{if } i = 2, \\ \Phi(\hat{c}_3 - \hat{\beta}_j) - \Phi(\hat{c}_3) - [\Phi(\hat{c}_2 - \hat{\beta}_j) - \Phi(\hat{c}_2)] & \text{if } i = 3, \\ \Phi(\hat{c}_4 - \hat{\beta}_j) - \Phi(\hat{c}_4) - [\Phi(\hat{c}_3 - \hat{\beta}_j) - \Phi(\hat{c}_3)] & \text{if } i = 4, \\ \Phi(\hat{c}_4) - \Phi(\hat{c}_4 - \hat{\beta}_j) & \text{if } i = 5, \end{cases} \tag{7}$$

where the evaluation is taken at $\mathbf{x} = \mathbf{0}$ (or at the left-out categories of the dummy variables) and $\hat{\beta}_j$ and $\hat{c}_i$, for $i = 1, ..., 4$, are design-consistent pseudo-maximum-likelihood estimates of the parameters.

Second, test statistics involving nonlinear functions of design-based estimators are simply more difficult to formulate. In this paper, we implement a Rao-Scott test statistic described in Graubard and Korn (1993), the formulation of which is given in the Appendix. It has the following form:

$$T_2 = \frac{n_{\text{March}} + n_{\text{June}}}{k_{\text{March}} + k_{\text{June}}} \frac{\hat{\Sigma}_{ij}}{M_i(\overline{\boldsymbol{\beta}}_j)} \sim \chi_1^2,$$

where n_q and k_q, for q = March, June, are respectively the sample size and the number of PSUs of survey period q,

$$M_i(\overline{\boldsymbol{\beta}}_j) = (n_{\text{March}} + n_{\text{June}}) \left[\text{Asy. V}\left(\left. \frac{\Delta P_i}{\Delta x_j} \right|_{\mathbf{x}=\mathbf{0},\text{March},\overline{\boldsymbol{\beta}}} \right) + \text{Asy. V}\left(\left. \frac{\Delta P_i}{\Delta x_j} \right|_{\mathbf{x}=\mathbf{0},\text{June},\overline{\boldsymbol{\beta}}} \right) \right],$$

in which $\overline{\boldsymbol{\beta}}$ is the vector of sample-weighted (simple random sampling) maximum likelihood estimates of $\boldsymbol{\beta}$, and $\hat{\Sigma}_{ij}$ is a consistent estimate of $\mathbf{G}_i(\boldsymbol{\beta}_j)\boldsymbol{\Gamma}_j\mathbf{G}_i(\boldsymbol{\beta}_j)'$ that is derived, say, from the jackknife replication method, with $\mathbf{G}_i(\boldsymbol{\beta}_j)$ as the 1×2 vector

$$\left[\frac{\partial}{\partial \beta_j} \left. \frac{\Delta P_i}{\Delta x_j} \right|_{\mathbf{x}=\mathbf{0},\text{March}} \quad \frac{\partial}{\partial \beta_j} \left. \frac{\Delta P_i}{\Delta x_j} \right|_{\mathbf{x}=\mathbf{0},\text{June}} \right],$$

which is evaluated at the true values of the infinite population parameters $\boldsymbol{\beta}$, and with $\boldsymbol{\Gamma}_j$ being the 2×2 asymptotic variance matrix of $\sqrt{k}\overline{\boldsymbol{\beta}}_j = \sqrt{k_{\text{March}} + k_{\text{June}}}\left[\overline{\beta}_{j,\text{March}} \quad \overline{\beta}_{j,\text{June}}\right]$.

III. Data Sets and Variables

The data used in this study are drawn from the March and June 1999 surveys of the Social Weather Stations (SWS). A private, non-stock, non-profit social research institution that was founded in 1985, the SWS conducts *social weather* surveys on a quarterly basis as a way of taking contemporary readings of Philippine economic and social conditions. Among the topics that the SWS regularly monitors are self-rated poverty, disposition on the quality of life, satisfaction regarding the performance of government officials and agencies, and the electoral prospects of various possible candidates.

The SWS surveys have a complex design. Specifically, they follow a stratified clustered random sampling scheme, in which the strata are the National Capital Region (NCR), the rest of Luzon, the Visayas, and Mindanao, and the primary sampling units (PSUs) are the municipalities in the case of NCR and the provinces in the case of the other areas. Each survey has a sample of 1200 respondents who are Filipinos of voting ages.

Table 1 presents the variables used in this study and their descriptive statistics for the regression sample in each of the surveys. It shows that the explanatory variables are sets of mutually exclusive binary variables that reflect the area and locale of the respondents' residence, their gender, age group, highest level of educational attainment, de facto conjugal status, socio-economic class, and religious affiliation. The statistics reported include the weighted means and design-based standard errors of the variables, as well as the magnitudes of what may be called a *design-effect* index. The weights used in the calculation of the means are the sample weights of the respondents (or the reciprocals of their selection probabilities), which means that the statistics are estimates of the finite-population means. The design-based standard errors are so-called because they take account of the complex survey design that generated the data on the variables. The design-effect index is simply the ratio of the design-based estimate of the variance (of a statistic) to the estimate of the variance that would have been obtained from a hypothetical survey of the same sample size, which is undertaken using simple random sampling without replacement. As such, it is a measure of the factor of magnitude by which test statistics based on the simple random sampling estimates of the variance can be off the mark.[6]

To provide a sense of the comparability or dissimilarity between the March and June surveys, Table 1 also reports the results of separate tests on the difference in population means of the variables. The outcomes indicate that only for three variables—the 25 to 29 year old age group, vocational school as the highest educational institution attended, and socioeconomic class D—are the estimated means statistically different between the two surveys. Hence, the inference may be made that, at least as far as our explanatory variables are concerned, the two surveys are more or less similar, and in effect validates the survey procedures of the SWS.

Finally, notice that the difference between the March and June estimates of the population means of the President's ratings is found to be statistically insignificant. This implies that concerns about declining presidential ratings at least as of June 1999 would have been unjustified.

IV. Estimation Results

Between March and June 1999, who changed their minds about Erap? This section presents and interprets the results of two methodological approaches that were implemented in this study to answer this very question. In the first method, (estimates of the finite) subpopulation means of the president's ratings and of his net satisfaction index (which is the statistic that the SWS generates and tracks) were generated for each survey period, in which the respondents were segregated according to their locational, demographic, and socioeconomic and cultural characteristics. These subpopulation means were then compared between the two survey periods to identify the respondent characteristics for which the president's ratings or his net satisfaction index had undergone a statistically significant change. In the second method, an ordered probit model of the president's ratings was estimated for each period, and the coefficients of the explanatory variables as well as their marginal effects on the probabilities of the ratings were then subjected to hypotheses tests.

---

[6] As may be gleaned from Table 1, the effect of the complex survey design on the estimated standard error of the variables can be rather large. In the March sample, for instance, it turns out that the design-based standard error of residence in the Visayas is as much as $\sqrt{42.6454} = 6.53$ times larger than the estimate of the simple-random-sampling-without-replacement standard error. In the June sample, on the other hand, the design-based standard errors of residence in the Visayas and Mindanao are, respectively, 7.25 and 8.15 times larger than their simple-random-sampling-without-replacement standard errors.

Table 2 presents the results of the first method. It indicates that, between March and June 1999, the subpopulation means of the president's ratings underwent a statistically significant change among residents of urban areas, particularly those in the Visayas and Mindanao, for persons whose highest level of education occurred in post-secondary vocational schools, and for individuals who were not affiliated with the Roman Catholic Church or Islam.[7] Except for persons whose religious leanings were classified as belonging to the "Others" category,[8] all of the statistically significant changes showed declines in the president's popularity.

Table 2 also compares the subpopulation means of the president's net satisfaction index for the same set of respondent characteristics. It discloses that the March and June subpopulation means of the index were statistically different for residents of urban areas in Mindanao, for rural inhabitants, and among individuals whose highest educational attainment was a post-secondary vocational program. Only for rural residents did the president's net satisfaction index show an increase.[9]

The results reported in Table 2 must be taken with caution, however, inasmuch as they may be tainted with omitted variable bias. The more robust results, which are those of the ordered probit model of the second method, are presented in Table 3.

Which factors had coefficient estimates that were significantly different from zero? Table 3 indicates that, for the March survey, they include residence in Mindanao (relative to the left-out category of residence in the National Capital Region (NCR)), having pursued a graduate degree (as opposed to having no formal schooling), belonging to socioeconomic class C (rather than class E, the poorest category), and being Muslim (rather than Roman Catholic). For the June survey, on the other hand, Table 3 reports that the respondent characteristics that had statistically significant coefficient estimates include residence in Luzon (outside NCR) and Mindanao and in urban areas of the two aforementioned areas, belonging to the 40 to 44, 45 to 49, 55 to 59, and 60 and over age groups (against being between 18 and 24 years old), having attended a post-college educational program, belonging to socioeconomic classes C and D, and being affiliated with a Protestant denomination, Islam, or other religious sect.

Our concern, however, is not so much to identify the regressors with statistically significant coefficient estimates, but to identify those factors whose coefficient estimates underwent a statistically significant change between the two survey periods. This is because it is the latter set that provides an indication of how the index function U^* changed over the two time periods. Hence, the last column of Table 3 reports the results of Wald tests on the hypothesis that

---

[7] For certain characteristics, such as not having had any formal schooling, having attended a post-college educational program, belonging to socioeconomic class AB, and being Muslim, the subpopulation means of the president's ratings could not be calculated because at least one of the strata had only one primary sampling unit. Moreover, the strata were thought to be already too few and too different for their PSUs to be merged in some way. Thus, we deemed it best to point out that the sample size of the SWS surveys may be too small for certain subpopulation analyses to be carried out.

[8] Included in this category are the *Iglesia ni Kristo* and the Philippine Independent Church.

[9] Although intended to track the same phenomenon (i.e., the rise and fall of the president's ratings), the two indicators show slightly different results because they emphasize different aspects of the ratings. Whereas the mean of the ratings indicates the location of the center of the frequency distribution, the mean of the net satisfaction index compares the relative fatness of the right and left tails of the distribution, completely disregarding the area taken up by the center.

the coefficient estimates for the same independent variable are equal between the two surveys. That column indicates that the coefficient estimates that underwent a statistically significant change are those of residence in an urban area in Mindanao, being between 35 and 39 years old or between 55 and 59 years old, and being affiliated with a Protestant denomination or Islam.

In qualitative-response models, though, the coefficient estimates by themselves do not indicate the marginal effect of a regressor on the dependent variable. Instead, these effects have to be calculated separately from the regression results (using (7)). Tables 4a and 4b present the results of these computations, in which the evaluations are taken at the left-out categories of the explanatory variables.[10] For the March survey, Table 4a reports that the statistically significant marginal effects are those of residence in Luzon (outside of NCR) and Mindanao and of class AB, the top socioeconomic category.[11] Specifically, the table indicates that, compared to the base person, a resident of Luzon with otherwise identical characteristics was 0.8 percentage points less likely to be dissatisfied with the president's performance as of March 1999, 1.0 percentage point less likely to be neither satisfied nor dissatisfied, 0.4 percentage points less likely to be satisfied, and 2.7 percentage points more likely to be very satisfied. In addition, the table suggests that a Mindanao resident with otherwise identical characteristics as the base person was 3.2 percentage points less likely to be dissatisfied with the president's performance, 4.8 percentage points less likely to be neither dissatisfied nor satisfied, 4.3 percentage points less likely to be satisfied, and 14.3 percentage points more likely to be very satisfied. On the other hand, in marked contrast to the marginal effects of geographic location, the results for socioeconomic class AB are that the probability that such a person was dissatisfied with the president's performance was 0.3 percentage points higher than that of the base person, the probability that she was neither dissatisfied nor satisfied was 0.3 percentage points higher, that she was satisfied was 0.1 percentage points higher, and that she was very satisfied was 0.8 percentage points lower.

In the case of the June survey, Table 4b reports that the statistically significant marginal effects are those of residence in the Visayas and of all the educational attainment variables. Specifically, Table 4b indicates that, compared to the base person, residents of the Visayas (with otherwise the same characteristics) had marginally lower probabilities of having been dissatisfied, having been neither dissatisfied nor satisfied, or having been satisfied, but they had a marginally higher probability of having been very satisfied. The table also suggests that (a) an individual who had at most an elementary education was 1.3 percentage points more likely to be dissatisfied with the president's performance than the base person, 1.4 percentage points more likely to be neither dissatisfied nor satisfied, 3.2 percentage points more likely to be satisfied, and 6.5 percentage points less likely to be very satisfied; (b) an individual who had at most a high school education was 1.9 percentage points more likely to be dissatisfied with the president's performance than the base person, 2.2 percentage points more likely to be neither dissatisfied nor satisfied, 4.4 percentage points more likely to be satisfied, and 9.4 percentage points less likely to be very satisfied; (c) an individual who had at most a vocational school education was 4.2 percentage points more likely to be dissatisfied with the president's

[10] The left-out categories combine to give the following profile of the base person against whom ratings of the other respondents are measured: a single, female, NCR resident, who is between 18 and 24 years old, has had no formal schooling, belongs to socioeconomic class E, and is Roman Catholic.

[11] The Rao-Scott test is implemented for Tables 4a and 4b using the null hypothesis that $\theta = \Delta P_i/\Delta x_j = 0$.

performance than the base person, 4.3 percentage points more likely to be neither dissatisfied nor satisfied, 6.5 percentage points more likely to be satisfied, and 2.7 percentage points less likely to be very satisfied; (d) an individual who had at most a college education was 4.7 percentage points more likely to be dissatisfied or to be neither dissatisfied nor satisfied with the president's performance, 6.7 percentage points more likely to be satisfied, and 18.4 percentage points less likely to be very satisfied; and (e) an individual who had at most a post-college education was 14.4 percentage points more likely to be dissatisfied with the president's performance than the base person, 10.4 percentage points more likely to be neither dissatisfied nor satisfied, 0.4 percentage points less likely to be satisfied, and 35.6 percentage points less likely to be very satisfied.

Which among the marginal effects were found to be statistically different between the two surveys? Table 4c, which reports the results of the second test statistic that was formulated for this study, indicates that the statistically different marginal effects were those of residence in Luzon (outside of NCR) and the Visayas as well as those of the educational attainment variables. As for the relative magnitudes of these statistically different marginal effects, it turns out that, with the exception of Luzon residents for whom the change in the probability of having been very satisfied with the president's performance (relative to that of the base person) was higher in June than in March, which in effect indicates a density function that shifted to the right, the other variables had marginal effects that suggest a density function that shifted to the left. This explains (a) the larger changes in the probabilities of having been very dissatisfied, having been dissatisfied, having been neither dissatisfied nor satisfied, or having been satisfied in June than in March among persons who had at most an elementary school education; (b) the larger changes in the probabilities of having been very dissatisfied, having been dissatisfied, or having been satisfied, but the smaller change in the probability of having been very dissatisfied in June than in March among persons who had at most either a high school or college education, (c) the smaller change in the probability of having been very dissatisfied in June than in March among persons who had at most a vocational education, and (d) the larger change in the probability of having been satisfied in June than in March among persons with at most a post-college education.

Hence, what the results in Table 4c imply is that, by June 1999, the common density function of U and U^* with respect to residence in the Visayas and the educational attainment variables had shifted to the left of where it had been in March 1999. In other words, disenchantment among residents of the Visayas and among persons who have had some education had started to seep in and spoil the Estrada magic, even though their effects on the gross indicators, i.e., the population means of the president's ratings and of the net satisfaction index, were still too small to be statistically significant.

V. Summary and Conclusion

In this paper, we explore the decline in President Estrada's satisfaction ratings in 1999. Using data from the March and June 1999 surveys of the SWS, we present and analyze the subpopulation means of President Estrada's ratings and of his net satisfaction index for certain respondent characteristics, in which the respondents were segregated according to their locational, demographic, and socioeconomic and cultural characteristics and the estimators used arre consistent with complex design of the surveys. In addition, we separately estimate a design-based ordered probit model of the president's ratings for each survey period and compare the parameter estimates as well as the marginal effects of the regressors on the probabilities of the ratings to identify the respondent characteristics that contributed to the net change in the president's satisfaction index.

Among our more significant findings are the following: (a) The effect of the complex survey design on the estimated standard error of the variables can be rather large. In the March sample, for instance, the design-based standard error of residence in the Visayas is 6.53 times larger than the estimate of the standard error that would have been obtained from a hypothetical survey of the same sample size, but which is based on a simple random sampling scheme without replacement; in the June sample, the design-based standard errors of residence in the Visayas and Mindanao are, respectively, 7.25 and 8.15 times larger. (b) The March and June 1999 data sets are more or less similar, at least as indicated by our estimates of the finite-population means. Only for three variables—the 25 to 29 year old age group, vocational school as the highest educational institution attended, and socioeconomic class D—were the population means statistically different between the two surveys. This is a validation of the survey procedures of the SWS. However, the sample size of the surveys—at 1200—although nationally representative, may be too small for certain subpopulation analyses. (c) Using parsimonious models, such as statistical two-way frequency tables of the president's ratings or his net satisfaction index and respondent characteristics (or subpopulation means of the ratings for categories of respondent characteristics) can lead to the wrong inferences being drawn. The results of our hypotheses tests on the subpopulation means of the president's ratings, for instance, suggest that March and June means were statistically different for residents of urban areas, particularly those in the Visayas and Mindanao, for persons whose highest level of education occurred in post-secondary vocational schools, and for individuals who were not affiliated with the Roman Catholic Church or Islam. In contrast, the statistically different marginal effects on the president's ratings that were observed from the ordered probit models are those of residence in Luzon (outside of NCR) and the Visayas as well as those of the educational attainment variables. (d) The coefficients estimates of the ordered probit models that turned out to be statistically different between the March and June 1999 surveys are those of residents of urban areas in Mindanao, of persons aged 35 to 39 years or 55 to 59 years, and of persons affiliated with Islam or with Christian denominations other than the Roman Catholic Church. In other words, these were the explanatory variables whose weights in the respondents' index functions changed between March and June 1999. (e) Although we find that the estimates of the population means of the president's ratings and of the net satisfaction index were not statistically different between the two survey periods, the marginal effects on the probabilities that were calculated from the ordered probit models reveal that, by June 1999, residents of the Visayas and individuals who have had some education had begun to become disenchanted with the Estrada administration. It was just that their incipient effects were just too small to be reflected in the gross indicators of the president's ratings.

Who changed their minds about Erap? The results of our ordered probit models indicate that the residents of Luzon (outside NCR) and of the Visayas did, as did people who have had some education. Moreover, our hypothesis tests indicate that, with the exception of Luzon residents who became more enamored with the president, residents of the Visayas as well as individuals who have had some schooling had become more disillusioned with the president.

Appendix

This section describes the formulation of the (second) test statistic that is used to test the hypothesis that the marginal effect of a regressor on a given ratings probability does not change over the two survey periods. It is a specific application of the Rao-Scott test statistic described in Graubard and Korn (1993).

We are interested in testing the following hypothesis with respect to the infinite population:

$$\theta = \left.\frac{\Delta P_i}{\Delta x_j}\right|_{\mathbf{x}=\mathbf{0},\text{March}} - \left.\frac{\Delta P_i}{\Delta x_j}\right|_{\mathbf{x}=\mathbf{0},\text{June}} = 0,$$

where $\Delta P_i/\Delta x_j|_{\mathbf{x}=\mathbf{0},\,q}$ is the marginal effect of the jth independent variable on the probability that rating i is chosen in survey period q, for $i = 1, 2, \ldots, 5$ and q = March, June, with the evaluation taken at $\mathbf{x} = \mathbf{0}$ or at the left-out categories of the dummy variables. It may be noted that θ can be expressed as a nonlinear function $g(\boldsymbol{\beta})$, where $\boldsymbol{\beta} = [\boldsymbol{\beta}_{\text{March}}, \boldsymbol{\beta}_{\text{June}}]$ is a vector of the infinite-population parameters of the two ordered probit models that are separately estimated (Please see the main text.).

If the March and June data sets were generated using simple random sampling procedures, a Wald test statistic can be formulated on the basis of the following argument: Since the marginal effects are nonlinear functions of estimators that are known to be asymptotically normally distributed, a first-order Taylor series approximation may be used to derive the result that the marginal effects are themselves asymptotically normally distributed with means equal to the right hand side expressions found in (7) in the main text, except that the cumulative standard normal distribution functions are evaluated at the true values of the parameters, and with asymptotic variance matrices of the form

$$\text{Asy.V}\left[\frac{\Delta P(U=i)}{\Delta \mathbf{x}}\right] = \frac{\partial}{\partial \boldsymbol{\beta}'}\left[\frac{\Delta P(U=i)}{\Delta \mathbf{x}}\right]\text{Asy.V}(\boldsymbol{\beta})\left\{\frac{\partial}{\partial \boldsymbol{\beta}'}\left[\frac{\Delta P(U=i)}{\Delta \mathbf{x}}\right]\right\}' \quad \text{for } i = 1,\ldots,5,$$

where, along the principal diagonal of the matrices of derivatives,

$$\begin{aligned}
\frac{\partial}{\partial \beta_j}\left[\frac{\Delta P(U=1)}{\Delta x_j}\right] &= -\varphi(c_1 - \beta_j) \\
\frac{\partial}{\partial \beta_j}\left[\frac{\Delta P(U=2)}{\Delta x_j}\right] &= [\varphi(c_1 - \beta_j) - \varphi(c_2 - \beta_j)] \\
\frac{\partial}{\partial \beta_j}\left[\frac{\Delta P(U=3)}{\Delta x_j}\right] &= [\varphi(c_2 - \beta_j) - \varphi(c_3 - \beta_j)] \\
\frac{\partial}{\partial \beta_j}\left[\frac{\Delta P(U=4)}{\Delta x_j}\right] &= [\varphi(c_3 - \beta_j) - \varphi(c_4 - \beta_j)] \\
\frac{\partial}{\partial \beta_j}\left[\frac{\Delta P(U=5)}{\Delta x_j}\right] &= \varphi(c_4 - \beta_j)
\end{aligned}$$

where φ(·) is the density function of the standard normal random variable, while, on the off-diagonal cells where $l \neq j$,

$$\frac{\partial}{\partial \beta_l}\left[\frac{\Delta P(U=i)}{\Delta x_j}\right] = 0 \quad \text{for } i = 1,\ldots,5.$$

Thus, under the assumption that the covariances of the March and June parameter estimates are zero and the result that the marginal effects are asymptotically normally distributed (with means and variance matrices as mentioned above), the hypothesis that the marginal effect of a regressor, x_j, on the probability that $U = i$ does not change between the March and June surveys may be evaluated using the following Wald test statistic:

$$W(\tilde{\boldsymbol{\beta}}) = \frac{(\tilde{\gamma}_{ij\,\text{March}} - \tilde{\gamma}_{ij\,\text{June}})^2}{\text{Asy. V}(\tilde{\gamma}_{ij\,\text{March}}) + \text{Asy. V}(\tilde{\gamma}_{ij\,\text{June}})} \sim \chi_1^2, \tag{A1}$$

where $\tilde{\gamma}_{ijq} = {}^{\Delta\tilde{P}(U=i)}\!\big/_{\Delta x_j}\Big|_q$, for i = 1, ..., 5, and q = March, June, is the marginal effect of regressor j on the ith rating probability, evaluated at $\tilde{\boldsymbol{\beta}}$, the maximum likelihood estimates of $\boldsymbol{\beta}$ using data generated by simple random sampling procedures.

Given that our data sets have a complex survey design, however, an alternative test statistic needs to be devised—one that hopefully is based on (A1), but in which the evaluation is performed using design unbiased estimators of $\boldsymbol{\beta}$. A prime candidate in this regard is $\overline{\boldsymbol{\beta}}$, the *weighted simple random sampling* maximum likelihood estimator of $\boldsymbol{\beta}$, in which the weights used are the sample weights of the observations. $\overline{\boldsymbol{\beta}}$ yields unbiased estimates of $\boldsymbol{\beta}$, but gives incorrect estimates of the variance matrix, since it does not account for the correlations between PSUs in each stratum.

Under certain regularity conditions, Graubard and Korn (1993) show that as the total number of PSUs k (in the two surveys) increases without bound, $\sqrt{k}(\overline{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma})$ and (under our specific null hypothesis) $nW_i(\overline{\boldsymbol{\beta}}_j) \xrightarrow{d} \chi_1^2$, where $\boldsymbol{\Gamma}$ is a variance matrix, n is the total sample size of the two surveys, and $nW_i(\overline{\boldsymbol{\beta}}_j) = c\Sigma_{ij} \big/ M_i(\overline{\boldsymbol{\beta}}_j)$, where $n/k \to c$, $M_i(\overline{\boldsymbol{\beta}}_j)$ is n times the denominator of (A1), except that it is evaluated at $\bar{\gamma}_{ij} = {}^{\Delta\overline{P}(U=i)}\!\big/_{\Delta x_j}$, and $\Sigma_{ij} = \mathbf{G}_i(\boldsymbol{\beta}_j)\boldsymbol{\Gamma}_j\mathbf{G}_i(\boldsymbol{\beta}_j)'$, in which $\mathbf{G}_i(\boldsymbol{\beta}_j)$ is the 1 × 2 vector

$$\left[\frac{\partial}{\partial \beta_j}\frac{\Delta P_i}{\Delta x_j}\bigg|_{\mathbf{x}=\mathbf{0},\text{March}} \quad \frac{\partial}{\partial \beta_j}\frac{\Delta P_i}{\Delta x_j}\bigg|_{\mathbf{x}=\mathbf{0},\text{June}}\right],$$

which is evaluated at the true values of the infinite population parameters $\boldsymbol{\beta}$, and $\boldsymbol{\Gamma}_j$ is the 2 × 2 asymptotic variance matrix of $\sqrt{k}\overline{\boldsymbol{\beta}}_j = \sqrt{k_{\text{March}} + k_{\text{June}}}\left[\overline{\beta}_{j,\text{March}} \quad \overline{\beta}_{j,\text{June}}\right]$. In turn, this allows us to formulate the following Rao-Scott test statistic:

$$T_2 = \frac{n_{\text{March}} + n_{\text{June}}}{k_{\text{March}} + k_{\text{June}}}\frac{\hat{\Sigma}_{ij}}{M_i(\overline{\boldsymbol{\beta}}_j)} \sim \chi_1^2,$$

where n_q and k_q, for q = March, June, are respectively the sample size and the number of PSUs of survey period q and $\hat{\Sigma}$ is a consistent estimate of Σ that is derived using, say, the jackknife replication method of parameter and variance estimation:

$$\hat{\beta}_j = \frac{1}{S}\sum_{s=1}^{S}\frac{1}{P_s}\sum_{p=1}^{P_s}\overline{\beta}_{-pj} \quad \text{and} \quad V(\hat{\beta}_j) = \sum_{s=1}^{S}\frac{P_s - 1}{P_s}\sum_{p=1}^{P_s}(\overline{\beta}_{-pj} - \hat{\beta}_j)^2,$$

where $\overline{\beta}_{-pj}$ is the sample weighted maximum likelihood estimate of β_j in which observations from the pth sampled PSU of stratum s are deleted and the sample weights of the remaining observations in the stratum are increased by a factor of $(P_s - 1)/P_s$.

References

Graubard, Barry I. and Edward L. Korn. 1993. "Hypothesis Testing with Complex Survey Data: The Use of Classical Quadratic Test Statistics with Particular Reference to Regression Problems." *Journal of the American Statistical Association* **88**: 629–641.

Greene, William H. 2000. *Econometric Analysis*. Fourth edition. New Jersey: Prentice-Hall.

Korn, Edward L. and Barry I. Graubard. 1999. *Analysis of Health Surveys*. New York: John Wiley and Sons.

Lee, Eun Sul, Ronald N. Forthofer, and Ronald J. Lorimer. 1989. *Analyzing Complex Survey Data*. California: Sage Publications.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

StataCorp. 1999. *Stata Reference Manual*. Release 6. Texas: Stata Press.

StataCorp. 1999. *Stata User's Guide*. Release 6. Texas: Stata Press.

Train, Kenneth. 1986. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand.* Cambridge, Massachusetts: MIT Press.