Causal Structure, Endogeneity, and the Missing Data Problem in Modeling the Impact of Information and Communication Technology (E-government) Use: Theoretical and Methodological Challenges

Hun Myoung Park
University Information Technology Services, Indiana University
kucc625@indiana.edu

Abstract

Despite the abundance of empirical research on the impact of information and communication technology, their relationship still remains partially answered because of conflicting results. Empirical research reports positive, negative, and negligible effects depending on data and methods employed. This puzzling circumstance results largely from the lack of rich data and sophisticated knowledge and skills. This paper reviews data analysis methods frequently used in the literature and then discusses key modeling issues, such as causal structure, endogeneity, and the missing data problem, which traditional methods rarely address. In order to deal with those issues, the propensity score matching, treatment effect model, and recursive bivariate probit model are suggested as alternatives. These methods do not replace but supplement traditional approaches. This paper concludes with the emphasis on careful examination of the characteristics of dependent variables and prudent consideration of the key modeling issues.

1. Introduction

As many citizens embrace information and communication technologies (ICT), scholars have turned their attention to the impact of ICT use on various aspects of society. One stream of research investigates how ICT use affects public service delivery, government reform, and democracy; another examines the impact on social capital and communities, digital inequality (digital divide), and cultural changes [4, 5, 11, 23, 24, 25].

Empirical studies on these topics are abundant in books, journals, reports, and conferences. Some studies suggest that ICT influences society positively, others report negative or negligible effects even when they use the similar data. This conflicting result is partially due to fast technology development and complexity of modern society.

Existing research employs various data analysis methods ranging from a simple descriptive method to a highly sophisticated model. Descriptive methods and content analysis such as Norris (2001) and West (2005) are rampant probably due to convenience and easy interpretation but difficult to draw conclusive inferences. The lack of rich data makes it unlikely that advanced econometric models, despite their methodological rigor and power, are widely applied in this field of study. This is a dilemma that many scholars are confronting.

Most quantitative empirical studies posit a simple univariate relationship between ICT use and society, and then employ common data analysis methods such as the t-test, ordinary least squares (OLS), and binary probit and logit model. These methods rarely address modeling issues associated with the causal relationship. This paper contends that the failure to consider those critical issues, such as the nature of causal structure, endogeneity, and the missing data problem, produces the conflicting empirical results.

Recently, e-government research moves beyond efficiency, transparency, and accountability and begins to address civic engagement, public outreach, and trust in government, and participatory democracy [8, 18, 28, 30, 34, 35]. The main issues in e-government research are not substantially different from those in general ICT research.

The goal of this paper is to improve theoretical and methodological rigor in the research on ICT use, in particular e-government use, by discussing key modeling issues and illustrating how alternative approaches deal with the problems involved. This paper argues that appropriate data analysis methods should be based on careful examination of the characteristics of individual dependent variables and prudent consideration of the major modeling issues.

In the next section, traditional data analysis methods are critically reviewed. Then discussed are limited dependent variables, causal structure, endogeneity, and the missing data problem. From section 4 through 6, the propensity score matching, treatment effect model, and recursive bivariate probit model are suggested as alternatives to deal with the modeling issues. Section 7 illustrates how individual methods estimate the treatment effect, compare their results, and summarize major findings followed by suggestions for future research.

2. Review of Traditional Approaches

Traditional approaches implicitly assume a simple unidirectional causality and random sampling; ICT use is determinant and samples are drawn randomly in experimental settings.[1] This section reviews three methods that are commonly used in the ICT literature.

2.1. T-Test (ANOVA)

The t-test compares the means of two groups. Analysis of variance (ANOVA) can compare more than two groups and also consider interaction effects among factors. These linear models have two underlying assumptions: normality and equal variance. Violation of the normality assumption calls for nonparametric methods, while unequal variance necessitates the approximation of the degrees of freedom.

This t-test (or ANOVA), largely due to its simplicity and easy interpretation, is frequently used to examine how ICT (e-government) use makes a difference. For example, Scott (2006) employs ANOVA to compare public involvement indices, which are developed from the content analysis of municipal government websites. The t-test can also analyze a binary response variable since the mean is interpreted as the proportion of the total.[2]

In an independent sample t-test, the two groups are implicitly assumed to have same characteristics except for the treatment. Thus, the effect, if any, is assumed to come from the treatment only. If covariates need to be controlled, analysis of covariance (ANCOVA), which bridges ANOVA and the linear regression model, is used instead.

2.2. Linear Regression Model

The classical linear regression model, called ordinary least squares (OLS), has been also widely used across disciplines. Jennings and Zeitner (2003), Uslaner (2004), and Welch and Pandey (2007) fit least squares estimators to analyze civic engagement, social capital, and website effectiveness and service quality of an organization. In OLS, the treatment effect is typically estimated using the least squares dummy variable model (LSDV). This is a fixed group effect model in a panel data setting [2, 13].[3] The functional form is

$$y_i = x_i'\beta + \delta d_i + \varepsilon_i$$

where δ is the coefficient of the dummy variable d; x is a vector of regressors (including the intercept) other than d; β is a vector of coefficients; and ε is disturbance (error term). The parameter estimate δ is the effect of ICT use (e-government use).

OLS is more informative than the t-test because it enables researchers to analyze the extent that individual independent variables affect the dependent variable. In LSDV, the treatment effect δ shifts the regression line vertically without influencing the slopes of regressors. The simple causal structure of OLS is easy to understand but vulnerable to critiques.

Among the key OLS assumptions is that regressors including d are not correlated with the disturbance ε. All regressors are exogenous or given from outside of the equation system. If some independent variables are endogenous or determined within an equation system, OLS estimates are biased and inefficient. The instrumental variable (IV) approach is often employed to deal with this endogeneity. However, it is not always easy to get good instrumental variables, which are highly related to problematic regressors and not related to the disturbance ε.

If researchers posit an equation system with endogeneity and simultaneity, the simultaneous equation model (SEM) and seemingly unrelated

[1] This paper focuses on quantitative (econometric) methods excluding such as content analysis, descriptive methods, nonparametric methods, and chi-squared test.

[2] In fact comparison of proportions should be based on the binomial probability distribution. When sample size is sufficiently large, however, both binomial and t distributions are approximated to the normal distribution, leaving no substantial difference in the effect size and interpretation between comparing proportions and means.

[3] This paper excludes advanced regression models associated with heteroscedasticity (random effect), autocorrelation, and time-series issues.

regression model (SUR) fit this causal structure [34]. As extensions of OLS, these econometric models focus more on cross-equation parameter estimates than the coefficient of an endogenous variable (e.g., e-government use). They also require rich data, which are not likely in many studies in social science.

In case of binary dependent variable, OLS (LSDV) produces biased and inefficient estimates with unrealistic prediction [14, 20]. Binary probit and logit models are used instead.

2.3. Binary Response Model

The binary response model (BRM) fits the nonlinear model of a binary dependent variable.[4] Binary probit and logit models are based on the standard normal and logistic probability distributions, respectively.[5]

$$P(y=1 \mid d, x) = \Phi(x\beta + \delta d)$$

$$P(y=1 \mid d, x) = \Lambda(x\beta + \delta d)$$

where Φ and Λ respectively indicate the cumulative standard normal and logistic distribution functions.

This is a fixed effect BRM. The treatment effect is not δ but defined as the discrete change, which means the difference in predicted probability between control and treated groups [14]. For instance, e-government users are $P(y=1 \mid x, x_k+\delta) - P(y=1 \mid x, x_k)$ percent more likely than nonusers to contact government officials, holding all other regressors at their baseline data points.

Since nationwide survey data tend to consist of many binary variables, BRM is often considered a promising econometric method. Bimber (2001, 2003) employs the binary logit model to examine how Internet use influences political participation such as attendance at a rally and financial contributions. Thomas and Streib (2003) applied this method to model use of Internet and government websites.

BRM in general requires a larger sample size than the t-test and OLS. Oftentimes, researchers try to fit sophisticated models with many parameters even when the sample size is small. This is the so called "small N, large parameter problem," which often becomes critical especially in nonlinear models. It is necessary to adjust the sample size and the number of parameters of a model in this circumstance.

3. Nature of Problems

The traditional methods discussed in the previous section are relatively simple and easy to apply. Meanwhile, they have strong underlying assumptions that are often violated in the real world. The phenomena to be explained tend to be more complex and problematic than expected. The failure to consider the nature of problems may end up with the fallacy of inferences.

3.1. Limited Dependent Variables

In general, it is not easy to measure social outcomes correctly. Imagine civic engagement, trust in government, and participatory democracy. Oftentimes, researchers get proxies of these variables or less informative, if not wrong, data (categorical rather than continuous variables). Most nationwide surveys such as the General Social Survey and the Pew Internet and American Life Project survey have many yes-no questions. It is often difficult and costly, if not impossible, to increase the sample size for the binary probit and logit models.

To make it worse, these dependent variables are often censored, truncated, and/or selected; they are frequently limited and thus problematic. A certain range of a dependent variable is not fully observed but transformed to a single value. Sample data are drawn from a subset of a large target population and thus some observations are not observed systemically. Reversely, the dependent variable is observed only for a selected nonrandom sample. Finally, individuals are not randomly assigned but they themselves decide whether or not to receive a treatment (self-selection).

These measurement issues and the characteristics of dependent variables often cast doubt on the relevance of traditional data analysis methods. OLS does not work well for categorical dependent variables, while BRM generally needs more observations than OLS. The presence of censoring, truncation, and/or selection calls for alternative approaches to fix the corresponding problems.[6] However, many studies do not appear to carefully examine the

[4] The discriminant analysis has been generally used to analyze categorical dependent variables. Largely due to the strong assumption of multivariate normality and difficulty in interpretation, probit and logit models (e.g., BRM, multinomial logit model, and conditional logit model) are currently preferred to this discriminant analysis.

[5] Despite this key difference in the probability distribution, these probit and logit models produce almost the same standardized effects of regressors [14].

[6] Since this paper focuses on estimating the treatment effect, censoring models (i.e., tobit model) and truncation models are skipped.

behavior of dependent variables. In most survey data, self-selection and the "missing data problem" are pervasive but oftentimes ignored in existing research.

3.2. Causal Structure

Many studies adopting the traditional methods envisage a unidirectional relationship between ICT use (e-government use) and society. That is, the former influences the latter in some way. This relationship, however, is often criticized as a misspecification because of the nebulous causal relationship between the two variables [4, 11]. They may be iterative and interactive in the virtuous circle [22]. Their relationship may be reciprocal and jointly determined simultaneously especially when a dependent variable is intensively associated with information exchange and deliberation.

The causal structure is not always clear and varies across specific dependent variables to be analyzed; the causation is direct and unidirectional in some relationships, but not in the other cases.

3.3. Endogeneity

In an econometric model, endogeneity occurs when some independent variables are correlated with the unobservable factors captured by the disturbance. Given a variety of factors intricated with each other in the real life, endogeneity is not exceptional but rather common.

The endogeneity results from the omission of relevant independent variables, measurement errors, and nonrandom sampling such as self-selection [2, 36]. ICT use is not always endogenous in a regression system but depends on type of the dependent variable.

In the presence of endogeneity, traditional models, say OLS, do not produce unbiased and efficient estimates. Hence, inferences from this model are not reliable. In case of endogeneity, the IV method (i.e., two-stage least squares) generally works if good IV variables are out there. Otherwise, more sophisticated approaches need to be considered.

3.4. Missing Data Problem

A perfectly controlled experiment can rule out biases through random assignment. But this randomized experiment tends to be costly, infeasible, and/or undesirable due to ethical and legal issues. In a semi- or pseudo-experimental design, more common in social science, the properties of the treated may be systematically different from those of control group. This pseudo-experimental design fails to remove potential spuriousness in a causal relationship and tends to under- or overestimate the treatment effect. The "missing data problem" oftentimes occurs in this circumstance.

If an individual receives the treatment, researchers cannot observe what his or her outcome would have been had he or she been assigned to the control group. This is the unobservable counterfactual outcome. Let y_1 denote the outcome of an individual who has used ICT applications (e.g., online forums and chat rooms of e-government) and y_0 be the one without e-government experiences. Let d denote a binary variable that indicates whether an individual visits government websites (d=1) or not (d=0). Since an individual cannot fall into both categories, either y_1 or y_0 is observed, but not both; a citizen is either an e-government user or nonuser, not both. This is called the "missing data problem."

$$y_i \equiv d_i y_{1i} + (1 - d_i) y_{0i}$$

Like endogeneity, the missing data problem comes from a couple of sources such as truncation and self-selection. The t-test and ANOVA estimates the average treatment effect (ATE), E(y_1 - y_0), which is the expected effect of treatment on participants who were randomly drawn from the population. However, ATE is biased if treatment (e.g., e-government use) is highly related to the dependent variable (e.g., attendance at a community meeting) in a semi-experimental setting. Suppose e-government use is positively related to citizens' willingness to contact government officials to express their policy opinions; an independent sample t-test may overestimate the impact of e-government use on this type of civic engagement.

What we need in this circumstance is the average treatment effect on the treated (ATET), E(y_1 - y_0 | d=1), which is the expected effect of treatment for those who received the treatment [36].[7] In order for reasonable comparison, we need to locate an individual in the control group "who looks like the treated one in every respect except for the treatment" [14]. However, this issue is seldom taken into account in existing e-government research.

In the following three sections, the propensity score matching, treatment effect model, and

[7] In a perfect experiment, ATE is identical to ATET since E(y_0 | d=1) is equal to E(y_0 | d=0).

recursive bivariate probit model are suggested as alternative methods to deal with the problems discussed so far.

4. Propensity Score Matching (PSM)

Since the seminal work of Rosenbaum and Rubin (1983), the propensity score matching (PSM) has been used in policy analysis and evaluations [1, 10, 19, 17]. PSM introduces one-dimensional propensity scores, predicted probabilities of falling into a treated group, to summarize multi-dimensional covariates [9, 15].

PSM is based on the "strongly ignorable treatment assignment" assumption that the treatment assignment d and outcomes of y_1 and y_0 are conditionally independent given covariates w [26]. "If treatment assignment is strongly ignorable, then adjustment for a balancing score $b(w)$ is sufficient to produce unbiased estimates of the average treatment effect" (44-45). The unobservable $E\{y_0 \mid b(w), d=1\}$ is drawn from the observable $E\{y_0 \mid b(w), d=0\}$ given propensity (balance) scores $b(w)$. By producing a control group whose distribution of covariates is similar to that of the treated group, PSM figures out this missing data problems [16].

The PSM method consists of four steps: 1) estimating propensity scores, 2) checking the balance of covariates, 3) matching (pair matching or subclassification), and 4) calculating average treatment effects and testing the null hypothesis [3, 10, 27].

Estimating Propensity Scores: The first step estimates propensity scores using either the binary logit or probit model. Propensity scores are predicted probabilities that citizens use ICT applications (e-government).

Checking the Balance of Covariates: the next step is to check if covariates of the control and treated groups are balanced by stratifying observations. Theorem 1 and 2 say that "treatment assignment and the observed covariates are conditionally independent given the propensity scores $p(w)$" or balancing scores $b(w)$ [26]. These theorems imply that if a subclass or stratum is homogeneous in propensity scores and some component of covariates, the treated and control groups of the stratum will have the same distribution of covariates and the components of covariates are expected to be balanced. If not balanced, the specification needs to be modified.

Pair matching or subclassicfication: When covariates are balanced in each stratum, observations are pair-matched or subclassified on the propensity scores [10, 27]. In the pair matching method, each treated observation is matched to a control observation with the closest propensity scores. Pair matching is performed by various methods such as nearest-neighbor matching, kernel matching, and Mahalanobis metric matching with or without replacement [3, 29]. In subclassification or stratification, observations are divided into subclasses or strata in which covariates are balanced across the treated and control observations. Control observations with a propensity score less than the minimum or larger than the maximum propensity score for treatment observations are discarded [10].

Calculating Average Treatment Effects and Testing the Null Hypothesis: PSM uses the paired sample t-test. In one-to-one pair matching, the difference, $\delta = \Sigma(y_1 - y_0)/N$, is considered the average treatment effect. The adjusted variance of the difference is computed using Bloom et al.'s (2002) formula. The degrees of freedom in one-to-one pair matching are the number of matched pairs minus one, N-1. The null hypothesis is that the mean difference between the treated and control groups is zero.

PSM does not assume any functional form or probability distribution but produces robust estimators of average effects when the missing data problem matters substantially. PSM, however, can tell only whether there is mean difference between the treated and control groups; it does not answer which covariate makes the difference, if any, and how [7]. Since it considers observed covariates only, PSM cannot control all biases that may or may not come from the unobserved [29]. The PSM method tends to be sensitive to data quality and how propensity scores are estimated.

5. Treatment Effect Model (TEM)

When individuals themselves decide whether or not they will receive the treatment (e.g., e-government use), the treatment effect model (TEM) can produce unbiased estimates of the effect by adjusting the selection bias. TEM is an econometric method that incorporates an endogeneous binary variable of the treatment into the regression equation [13, 21].

$$y_i = x_i'\beta + \delta d_i + u_i$$

$$d_i^* = w_i'\gamma + v_i,\ d_i = 1 \text{ if } d_i^* > 0, \text{ 0 otherwise}$$

where y is a continuous (interval) dependent variable; x and w are the regressor vectors; d is a binary variable for treatment; β, γ, and δ are parameter vectors; and u and v are normally distributed disturbances with a correlation ρ.

The selection equation, the second equation above, is used to compute the hazard function. Unlike Heckman's selection model, TEM uses both e-governmnet users and nonusers in analysis.[8] Unlike LSDV, the treatment effect, E(y| d=1, x, w)- E(y | d=0, x, w), is not δ but depends on ρ, σ_u, $\Phi(w\gamma)$, and □($w\gamma$).[9] Treatment effect is

$$\delta + \rho\sigma_u\left(\frac{\phi_i}{\Phi_i(1-\Phi_i)}\right)$$

The second term of the above indicates the influence of self-selection to be adjusted. When ρ=0 (no significant self-selection), the impact of e-government use becomes δ as in LSDV.

This TEM assumes that the control and treated groups have the same slopes but different intercepts. If researchers posit two regimes that have different slopes and intercepts, the (endogenous) switching regression model will be the case [13, 21].

6. Recursive Bivariate Probit Model

The recursive bivariate probit model (RBPM) is proposed by Maddala (1983) and is developed further by Greene (1998, 2003). RBPM assumes that two dependent variables y and d are jointly determined with a weak causation. Its functional form is

$$y^* = x_1^{'}\beta_1 + \delta d + \varepsilon_1,\quad y = 1 \text{ if } y^* > 0$$
$$d^* = x_2^{'}\beta_2 + \varepsilon_2,\qquad d = 1 \text{ if } d^* > 0$$

where y is a binary dependent variable of interest in equation 1, d is a binary dependent variable of equation 2 that is included in the first equation as an endogenous independent variable, and x_1 and x_2 are the regressor vectors of two regression equations. The disturbance ε_1 and ε_2 are normally distributed with a correlation of ρ. Note that both dependent and endogenous independent variables are binary.

This equation system is identified if disturbances are independent or there is at least one exogenous variable in x_2 that is not included in x_1 [21]. Interestingly, the endogenous nature of d in the first equation can be ignored in formulating the likelihood function [13]; d can be used as an endogenous variable in the first equation as if there is no simultaneity problem because two dependent variables are jointly determined [12, 13].

The null hypothesis is that the disturbances ε_1 and ε_2 are not correlated, ρ=0. The correlation coefficient between the disturbances measures the effect after the influence of the endogenous variable d is accounted for in the first equation [13]. If the null hypothesis is not rejected (no correlation), the two equations may be estimated separately by either the binary logit or binary probit model.

Because RBPM is nonlinear, effects of individual parameter estimates should be interpreted with special caution. The impact of an independent variable on predicted probabilities is not constant, but depends on the values of that variable and other independent variables. [10] Marginal effects and discrete changes are often used to evaluate the effects of independent variables in logit and probit models [13, 20]. The marginal effect of a continuous variable consists of direct and/or indirect effects [12, 14].[11]

The discrete change of a binary variable is the change in conditional predicted probabilities when the variable changes from 0 to 1, holding other variables at their reference points. Likewise, the discrete change of the endogenous binary variable d is the difference in conditional predicted probability between e-government users and nonusers when other independent variables are held at their reference points.

7. Illustration

This section illustrates how each model estimates the impact of Internet and e-government use on civic engagement using the 2004 Post-election Internet Tracking Survey of the Pew Internet and American Life Project and the longitudinal survey data of the American National Election Studies (ANES). Because dependent variables available in survey data sets are binary, TEM is skipped here.

---

[8] Heckman's two-step selection model employs the inverse Mills ratio to deal with sample selection bias and uses observed (selected) data only. Unlike TEM, this "Heckit" is not primarily designed to estimate the treatment effect.

[9] □ is the standard normal probability density function.

[10] Unlike BRM and RBPM, linear models such as OLS and TEM report the impact (slope) of an independent variable, which is constant regardless of values of other variables.

[11] If an independent variable appears only in the first equation, the variable has only a direct impact. If an independent variable is included in the second equation, the variable has an indirect effect only. An independent variable appearing in both equations has both direct and indirect effects.

7.1 Data and Specification

The first data set is the 2004 Post-election Internet Tracking Survey, in which a nationally representative sample is drawn from adults living in continental U.S. households using the standard list-assisted random digit dialing method. The data set includes a total of 2,146 observations, excluding those with missing values in age.

Dependent variables are whether citizens sent email about voting or the campaign and whether they attended a rally during the election campaign. Treatment is whether citizens looked for information from a local, state, or federal government websites. Independent variables include political knowledge, mobilization, partisanship, family income, education, gender, race, age. For the e-government use equation, Internet experience, online use intensity, and broadband use were also used.

The ANES survey data of the election years of 1996, 1998, 2000, and 2004 were extracted from the cumulative data file from 1948 through 2004. The sample size is 6,014. The binary variable for treatment is whether citizens have used the Internet for political information.

Dependent variables are discussion about politics and financial contributions. The former asks if he or she talked about politics with family and friends at least once in the past week. The question of financial contributions reads "Did you give money to an individual candidate or political party running for public office?" Independent variables include political interest, knowledge, efficacy, mobilization, partisanship, social capital, trust in government, and demographics such as personal income, college education, age, gender, race. In the selection equation of the probit model, the demographics, political interest, knowledge, social capital, media exposure, and dummy variables for three election years (i.e., 1996, 1998, and 2004) are used as covariates.

7.2 Estimating Treatment Effects

Table 1 presents the treatment effects of e-government use on sending emails and attending a campaign rally. The independent sample t-test is compared to PSM, while BRM and RBPM are comparable to each other. The t-test and PSM produce the average effects (or mean difference), whereas RBM and RBPM report discrete changes as the effect of e-government use. This impact is the difference in (conditional) predicted probability between e-government users and nonusers. These discrete changes were calculated at the reference points of the means (or medians) of continuous independent variables and 1 for binary variables.

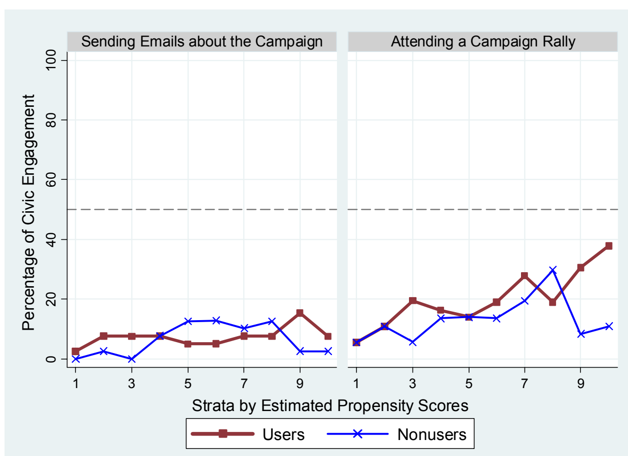
The independent sample t-test suggests that e-government users are 17.1 (=29.4-12.3) and 6.6 (=13.0-6.3) percent more likely than nonusers to send emails and attend a rally. However, PSM reports much smaller effects of 9.8 (=22.2-12.4) and 2.2 (=8.4-6.3) percent for both engagement activities. Figure 1 illustrates how average treatment effects change when predicted probabilities (propensity scores) of e-government use increase. The gap between users and nonusers in a stratum indicates the impact of e-government use in the stratum Attendance at a rally appears to be positively related to the likelihood of e-government use (see the right plot in Figure 1).

Table 1. Treatment Effects (E-gov. Use)

Method	Email	Rally
T-test	*17.1%*	6.6%
	(1,243)	(1,320)
PSM (Pair)	9.8%	2.2%
	(509)	(558)
BRM (Probit)	*14.1%*	3.3%
	(1,030)	(1,090)
RBPM	15.3%	*3.3%*
	(931)	*(974)*

Data source: The Pew Internet and American Life Project.
* The number of observations (or pairs) in parentheses.

Figure 1. Average Effect of E-gov. Use



The RBPM of sending emails has significant correlation between disturbances, which implies a substantial endogeneity in the equation system. This correlation measures the effect after the influence of government use on sending emails is accounted for. Therefore, BRM may not produce reliable results. The impact of e-government use on this deliberative engagement, the discrete change in conditional predicted probability, is 15.3 percent at the reference points, which is larger than the average treatment effect of 9.8 percent.[12] This result means that e-government users are 15.3 percent more likely than nonusers to send emails about voting or the campaign. RBPM allows researchers to analyze direct and indirect effects of an independent variable. For example, family income negatively influences sending emails as a whole despite its positive indirect impact through facilitating e-government use. In the model of attendance at a rally, e-government use does not appear endogenous but exogenous. BRM reports a discrete change of 3.3 percent, which is similar to the average effect that PSM produces.

Table 2. Treatment Effects (Internet Use)

Method	Discussion	Donation
T-test	*21.0%*	6.3%
	(5,419)	(5,425)
PSM (Pair)	10.1%	4.4%
	(1,091)	(1,090)
BRM (Probit)	*9.9%*	5.4%
	(4,956)	(4,959)
RBPM	8.3%	*5.2%*
	(4,956)	*(4,959)*

Data source: The American National Election Studies.
* The number of observations (or pairs) in parentheses.

Figure 2. Discrete Change of Internet Use

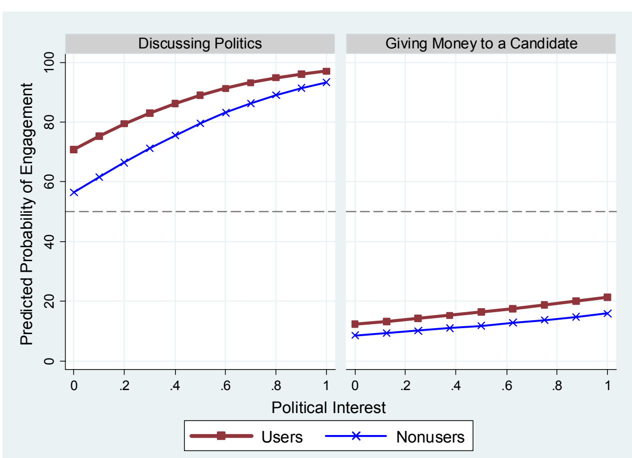


Table 2 presents the treatment effects of Internet use. The large sample size of the ANES data appears to make estimation easier and more reliable. The independent sample t-test reports the large average treatment effect of 21.0 percent on discussion about politics, which is almost double the average effect of PSM. Internet users are 10.1 percent more likely to discuss politics than nonusers. The t-test appears to overestimate the average treatment effect. Unlike this deliberative engagement, action-oriented financial contributions are not significantly influenced by e-government use. The t-test and PSM produce similar results of 6 and 4 percent, respectively.

Internet use is endogenous in the RBPM of discussion on politics. BRM does not capture this endogeneity and may produce misleading results. The discrete change of Internet use is 8.3 percent, which is similar to the average effect of PSM. The large indirect effect of personal income indicates that the income positively influences discussion on politics largely by facilitating Internet use rather than by boosting this engagement directly. In the model of financial contributions, there is neither significant correlation between disturbances nor substantial difference in discrete changes between BRM and RBPM. Figure 2 depicts predicted probabilities of the two engagements as citizens' political interest changes, holding other variables at their reference points. The vertical distance of users and nonusers' curves at a level of political interest indicates the impact of Internet use at the interest level.

7.3 Findings

There are several findings. First, PSM tends to report robust estimation of the treatment effect at the expense of unmatched data. PSM tends to use relatively a small portion of total observations because of exclusion of unmatched observations. In Table 2, about 11 hundred pairs (22 hundred observations) were used to estimate the average effect.[13] The overestimation of the t-test appears large in deliberative engagement and small in action-oriented engagement. This finding suggests that the missing data problem varies across dependent variables.

Second, BRM tends to overestimate or underestimate the treatment effect when significant endogeneity is involved in the

[12] This online engagement was measured only by Internet users and may thus involve other problems that RBPM cannot not handle properly. In addition, this RBPM may suffer from the "small N, large parameter problem."

[13] Linear and nonlinear models (i.e., OLS, TEM, BRM, and RBPM) employ pairwise deletion for missing values and thus use more observations in analysis than PSM.

model.[14] Otherwise, BRM and RBPM produce similar results. In the RBPM of attendance at a rally, the correlation of disturbance is not statistically discernable from zero; therefore, the impact of e-government use is direct and exogenous. The RBPM of discussion on politics involve significant endogeneity and the discrete change is smaller than that of BRM.

Third, RBPM provides an effective way of interpreting results in the presence of significant endogeneity. Its two equation system enables researchers to analyze direct and indirect effects of individual regressors. For instance, age may have a negative indirect effect in the selection (e-government use) equation and but a positive direct effect in the engagement equation. These conflicting direct and indirect effects may cancel each other out and thus the overall impact of age may become negligible. In this case, BRM would mistakenly report a significant positive (direct) effect only.

Finally, citizens are more likely to use ICT applications as a tool of information acquisition and exchange in deliberative engagement. ICT use in action-oriented engagement appears to be rather instrumental in a sense that convenience and efficiency gain are of primary interest of users and providers.

8. Conclusion

This paper critically reviews traditional methods that are frequently employed in the existing literature, discusses major modeling issues involved, and then suggests alternative approaches to deal with the key problems. However, the suggested methods are not necessarily better than the traditional ones; it depends on characteristics of dependent variables of interest and the nature of the causal relationship. If data are obtained from well designed experimental setting, traditional methods work well. Otherwise, researchers need to take advantage of the methodological rigor of alternative approaches.

Therefore, researchers always have to examine the data generation process carefully and understand the behavior of dependent variables. And then they should build a model that reflects a causal structure with endogeneity, the missing data problem, and other issues considered. If a dependent variable is continuous and involves self-selection, TEM and PSM may produce the unbiased treatment effect. If endogeneity resulting from self-selection and other sources matters in a model of a binary variable, RBPM and PSM may report unbiased estimates. Otherwise, T-test, LSDV, and BRM (for a binary dependent variable) will produce the reliable treatment effect. When censoring or truncation also matter, modeling will become more complicated and challenging.

Some studies simply report goodness-of-fit, parameter estimates, and statistical significance without substantive interpretation. This practice often fails to persuade audiences especially when presenting nonlinear models. The marginal effect for continuous independent variables and the discrete change for binary variables in probit and logit models are very useful to interpret results substantively [20].

Finally, ICT use (e-government use) influences various aspects of society in different ways. Its impact on deliberative engagement, for example, is different from the effect on attendance at a campaign rally and contact with government officials. Some activities of interest may involve endogeneity, while others may not.[15] Likewise, individual ICT applications (e.g., online forums, online consultations, and chat rooms of e-government) should be differentiated to illuminate their distinct implications for society [5, 10].

9. References

[1] Angrist, J.D., "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 66(2), 1998, 249-288.

[2] Baltagi, B.H., *Econometric Analysis of Panel Data*, Wiley, John & Sons, 2001.

[3] Becker, S.O., and A. Ichino, "Estimation of Average Treatment Effects Based on Propensity Scores," *STATA Journal*, 2(4), 2002, 358-377.

[4] Bimber, B., "Information and Political Engagement in America: The Search for Effects of Information Technology at the Individual Level," *Political Research Quarterly*, 54(1), 2001, 53-67.

[5] Bimber, B., *Information and American Democracy: Technology in the Evolution of Political Power*, Cambridge University Press, New York, 2003.

[14] Hausman specification test compares BRM and RBPM, but is not always possible since the variance difference matrix often becomes negative definite [14].

[15] Verba, Scholzman, and Brady (1995) differentiate political activities according to the capacity to convey information (or messages), strength of pressures, and required resources (43-48).

[6] Bloom, H.S., C. Michalopoulos, C.J. Hill, and Y. Lei, *Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?* MDRC Working Papers on Research Methodology, Manpower Demonstration Research Corporation, 2002.

[7] Bryson, A., R. Dorsett, and S. Purdon, *The Use of Propensity Score Matching in the Evaluation of Active Labour Market Policies*, Policy Studies Institute and National Centre for Social Research, 2002.

[8] Chadwick, A., and C. May, "Interaction between States and Citizens in the Age of the Internet: "e-Government" in the United States, Britain, and the European Union," *Governance*, 16(2), 2003, 271-300.

[9] D'Agostino, R.B., and D.B. Rubin, "Estimating and Using Propensity Scores with Partially Missing Data," *Journal of the American Statistical Association*, 95(451), 2000, 749-759.

[10] Dehejia, R.H., and S. Wahba, "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94(448), 1999, 1053-1062.

[11] DiMaggio, P., E. Hargittai, W.R. Neuman, and J.P. Robinson, "Social Implications of the Internet," *Annual Review of Sociology*, 27(1), 2001, 307-336.

[12] Greene, W.H., "Gender Economics Courses in Liberal Arts Colleges: Further Results," *Journal of Economic Education*, 29(4), 1998, 291-300.

[13] Greene, W.H., *Econometric Analysis*, Prentice Hall, Upper Saddle River, NJ, 2003.

[14] Greene, W.H., *LIMDEP Version 9.0 Econometric Modeling Guide*, Econometric Software, Plainview, NY, 2007.

[15] Hahn, J. "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 66(2), 1998, 315-331.

[16] Heckman, J.J., and S. Navarro-Lozano, *Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models*, National Bureau of Economic Research, 2003.

[17] Heckman, J.J., H. Ichimura, and P.E. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64(4), 1997, 605-654.

[18] Jennings, M.K., and V. Zeitner, "Internet Use and Civic Engagement," *Public Opinion Quarterly*, 67(3), 2003, 311-334.

[19] LaLonde, R.J., "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4), 1986, 604-620.

[20] Long, J.S., *Regression Models for Categorical and Limited Dependent Variables. Advanced Quantitative Techniques in the Social Sciences*, Sage Publications, 1997.

[21] Maddala, G.S., *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York, 1983.

[22] Norris, P., *A Virtuous Circle: Political Communication in Post-Industrial Democracies*, Cambridge University Press, New York , 2000.

[23] Norris, P., *The Digital Divide: Civic Engagement, Information Poverty and the Internet Worldwide*, Cambridge University Press, New York, 2001.

[24] Rheingold, H., *The Virtual Community*, Addison-Wesley Publishing, 1993.

[25] Robbin, A., C. Courtright, and L. Davis, "ICTs and Political Life," *Annual Review of Information Science and Technology*, 38, 2004, 411-481.

[26] Rosenbaum, P.R., and D.B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), 1983, 41-55.

[27] Rosenbaum, P.R., and D.B. Rubin, "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *Journal of the American Statistical Association*, 79(387), 1984, 516-524.

[28] Scott, J.K., "'E' the People: Do U.S. Municipal Government Web Sites Support Public Involvement?" *Public Administration Review*, 66(3), 2006, 341-353.

[29] Shadish, W.R., T.D. Cook, and D.T. Campbell, *Experimental and Quasi-experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston, MA, 2002.

[30] Thomas, J.C., and G. Streib, "The New Face of Government: Citizen-Initiated Contacts in the Era of E-Government." *Journal of Public Administration Research and Theory*, 13(1), 2003, 83-102.

[31] Uslaner, E.M., "Trust, Civic Engagement, and the Internet," *Political Communication,* 21(2), 2004, 223-242.

[32] Verba, S., K. Schlozman, and H.E. Brade, *Voice and Equality: Civic Voluntarism in American Politics*, Harvard University Press, Cambridge, MA, 1995.

[33] Welch, E.W., and S. Pandey, "Multiple Measures of Website Effectiveness and Their Association with Service Quality in Health and Human Service Agencies," *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, 2007.

[34] Welch, E.W., C.C. Hinnant, and M.J. Moon, "Linking Citizen Satisfaction with E-Government and Trust in Government," *Journal of Public Administration Research and Theory*, 15(3), 2005, 371-391.

[35] West, D.M., *Digital Government: Technology and Public Sector Performance*, Princeton University Press, Princeton, NJ, 2005.

[36] Wooldridge, J.M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA, 2002.