選舉研究
第二十二卷第一期（104/05），pp.01-33

Endogenous Regressors in Nonlinear Probability Models: A Generalized Structural Equation Modeling Approach*

Chi Huang**

Abstract

Endogeneity of explanatory variables is a common problem in many areas of social sciences. Ironically, there seems to be a gap between being aware of the problem and knowing how best to handle it. The problem is exacerbated when the outcome variable of interest is categorical and thus non-linear probability models are involved. The study fills the gap by first distinguishing two main sources of endogeneity, including unmeasured confounders ("latent factors") and measured but omitted causes ("endogenous mediators"), and then proposing an integrated approach to confront the two problems simultaneously. This strategy generalizes structural equation models to categorical outcome by including a shared latent factor between correlated error terms to tackle unobserved confounders, on the one hand, and extending mediation analysis to deal with potentially endogenous discrete mediators, on the other hand. For illustrative purpose, this proposed modeling strategy is presented with an example of heated debates in economic voting literature concerning the possible endogeneity of voters' economic perceptions.

Keywords: endogeneity, nonlinear probability models, mediation analysis, coefficient-rescaling problem, economic voting

---

* An earlier version of this paper was presented at the "2014 International Conference on Taiwan's Election and Democratization Study: Perspective on Citizens' Participation" held at National Chengchi University on March 22, 2014. The author would like to thank comments and suggestions from the discussant Professor John Fu-sheng Hsieh as well as two anonymous reviewers. The author alone is responsible for the arguments made in this article.

** Chair Professor, Department of Political Science and Senior Research Fellow, Election Study Center, National Chengchi University. E-mail: chihuang@nccu.edu.tw.

Empirical researchers are often interested in evaluating the effect of an explanatory variable on a particular outcome. Regression analysis is perhaps the most popular tool for this purpose. However, if there are factors not included in the analysis are correlated with this included explanatory variable and the error term (and hence with the outcome variable), the standard exogeneity (i.e., conditional independence) assumption for consistency is violated. In these cases, the estimated impact on outcome is confounded by the omitted factors. Such an explanatory variable is often called "endogenous regressor."

Endogeneity of explanatory variables is a common problem in many areas of empirical political science. Ironically, however, there seems to be a gap between being aware of the problem and knowing how best to handle it. As a result, endogeneity sometimes becomes a term shrouded in mystery and controversies. The problem is exacerbated when the outcome variable of interest is categorical and thus non-linear probability models are involved.

The purpose of this paper is to narrow the gap by pulling together related methodological developments scattered widely in different fields and disciplines. I try to clarify the nature of endogeneity by distinguishing its two main sources, unmeasured confounders ("latent factors") and measured but omitted causes ("endogenous mediators"), each of which calls for different solutions. After briefly reviewing the methods developed in the linear equations in the next section, I proceed in the third section to discuss additional severe challenges in nonlinear probability models for categorical outcome variables. The fourth section proposes an integrated parametric approach to the thorny problems of endogenous regressor in nonlinear models. This approach uses structural equation models (SEM) to include mediation analysis to deal with potentially endogenous discrete mediators, on the one hand, and allow correlated error terms to tackle unobserved confounders, on the other hand. For illustrative purpose, this proposed modeling strategy is presented with an example of heated debates in economic voting literature concerning the possible endogeneity of voters' economic perceptions. The last section concludes.

I. Ubiquitous Endogeneity Problem in Social Science Research

Endogeneity can occur as a consequence of simultaneous determination, sample selection, errors in variables, or the omission of relevant attributes that are correlated with the observed ones (Jackson 2008). Most of these problems lead to the confounding between a focused explanatory variable and the outcome variable. Before addressing this issue in nonlinear probability models, let us have a quick review of endogeneity in the more familiar linear models.

1. Confounded by Unobservables

The statistical symptom of endogeneity due to unmeasured confounders in regression models is the right-hand-side (RHS) explanatory variable being correlated with the error term. Under this type of model misspecification, the estimated coefficients of a linear or nonlinear model are biased and inconsistent. The key to solve the problem then lies in breaking the correlation between the endogenous regressor and the error term. In experiments, random assignment of the treatment is the gold standard of assuring exogeniety between treatment and potential outcomes. In observational studies, which social sciences including political science rely heavily upon, endogeneity problem of explanatory variables becomes ubiquitous.

The standard approach to addressing endogeneity attributed to unobserved (or latent) confounders in linear models is called instrumental variables (IV) method, first pioneered by econometricians then became a staple of tool kit in other fields (Bollen 2012; Sovey and Green 2011). Instrumental variable is either an observed variable served as the proxy of the endogenous regressor, or is constructed by projecting the "contaminated" variable onto a space so as to satisfy the "exclusion condition," i.e., IV has to be correlated with the endogenous regressor but, at the same time, orthogonal to the error term and hence breaks the dependence. Two-stage least squares (2SLS) is a familiar technique to achieve this goal. 2SLS is so called because it typically consists of two linear regression equations run in sequence. The first stage is an auxiliary regression fitted in order to construct the predicted value of the targeted endogenous variable. Then in the second stage this constructed predicted value, which is uncorrected with the error term by construction, is inserted into the structural equation of main interest to substitute the endogenous explanatory variable. Hence this strategy is also called two-stage predictor substitution (2SPS) (Terza, Basu, and Rathouz 2008). Despite the two equations involved in the process of analysis, 2SLS/2SPS is fundamentally a single-equation method focusing on the consistent estimation of the second-stage structural equation.

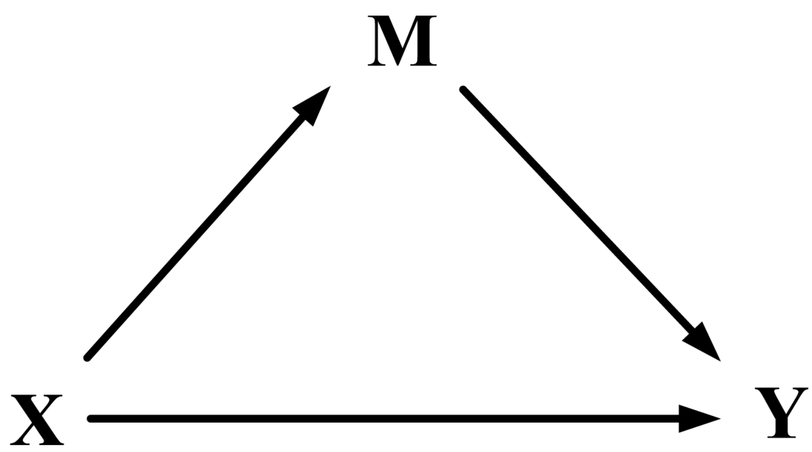
2. "Confounded" by Observables

In the political science literature, the term endogeneity is sometimes used loosely to mean that a RHS variable is not exogenous simply because it is affected by another *observed* variable.

This use most often occurs when critics challenge the theoretical position of an explanatory variable in a model. One example in the literature of economic voting is the argument that economic perception is endogenous because it may be structured by other *measured* political factors (see, for example, Anderson, Mendes, and Tverdova 2004; Wlezien, Franklin, and Twiggs 1997). Here the central focus is actually the theoretical status of the targeted explanatory variable, say M, in a causal system. That is, if X affects M and M in turn influences Y, then the effect of X on Y is (partially) mediated by M. I prefer calling such an explanatory variable a "mediator" so as to distinguish it from the confounding of latent variables discussed in the last section. Since X is observed, the best way to handle M is not just to question its exogeneity but to model X-M-Y relationships so as to test competing theories concerning M. This is exactly the subject of "mediation analysis" popularized by psychologists Baron and Kenny (1986).

Mediation analysis is widely used in the field of psychology and penetrated into social and biomedical sciences (see, for example, Hayes 2013; MacKinnon 2008). It investigates the mechanisms that underlie an observed relationship between a primary independent variable[1] and an outcome variable by examining how they relate to a third intermediate variable. Rather than hypothesizing only a direct causal relationship between the independent variable X and the dependent variable Y, a mediational model hypothesizes that the independent variable X affects the mediator variable M, which in turn affects the outcome variable Y. The simplest form of a mediation model can be illustrated as Figure 1, often dubbed the "golden triangle." The mediator M sitting on top of the triangle serves to illuminate the mechanisms through which X influences Y. It plays a pivotal role in understanding how the underlying process links X to Y and thus is no less, if not more, important than X. Therefore, the goal of mediation analysis goes beyond direct X-Y relationship by first testing the existence of X-M-Y relationship, and once established, estimating the extent to which the causal variable X influences the outcome Y through one or more mediator variables M, called mediation (or indirect) effect. The effect of X on Y not mediated by M is called the direct effect.

[1] In experimental designs, the independent variable X of primary interest is often called "treatment variable." In this paper, we focus exclusively on observational studies and assume that X is conditionally independent given the covariates (Pearl 2009).



Source: Hayes (2013, 91).

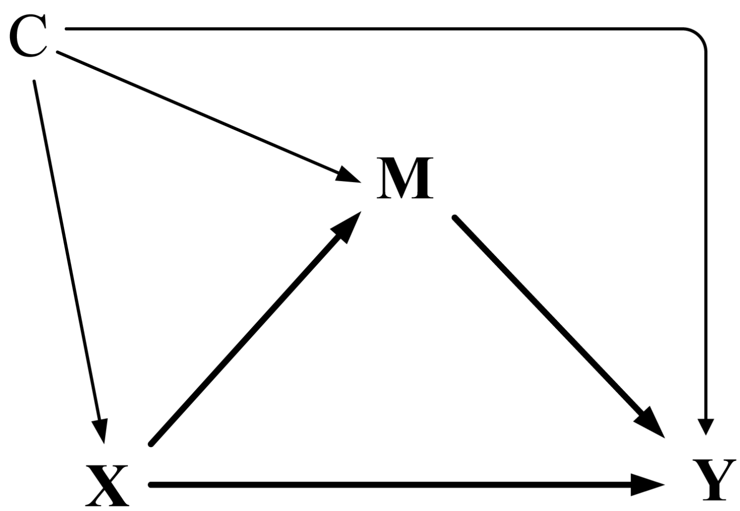
Figure 1 A Diagram of the Simple Mediation Model without Covariates

When both the mediator and the outcome variables are continuous, standard mediation analysis is usually conducted in the regression-based path-analytic framework. For a simple mediation model with covariates in Figure 2, it is typically represented by two linear regression equations, with C stands for a vector of all other covariates:

$$(1.1)\quad M_i = \alpha_1 + \gamma_1 X_i + \boldsymbol{C}_i \boldsymbol{\gamma}_2 + \epsilon_{1i}$$

$$(1.2)\quad Y_i = \alpha_2 + \beta_2 M_i + \gamma_3 X_i + \boldsymbol{C}_i \boldsymbol{\gamma}_4 + \epsilon_{2i}$$

In this setup, the standard method is to estimate the mediation effect using the *product of coefficients* of X and M, $\hat{\gamma}_1\hat{\beta}_2$, while estimate the direct effect by $\hat{\gamma}_3$, where all the coefficients are obtained by separately fitting ordinary least squares (OLS) regressions based on the two equations above. Alternatively, one can also estimate mediation effect by comparing the coefficient of X from a linear regression excluding the mediator M with the corresponding coefficient from a linear regression including M. The latter method is called "*difference in coefficients*."



Source: By the Author.

Figure 2 A Diagram of the Simple Mediation Model with Covariates

Although often mentioned only in passing at best, it should be emphasized here that this path-analytic approach applies only if the error terms of these two equations are uncorrelated. If ϵ_{1i} and ϵ_{2i} are correlated, however, a separate OLS estimation of the M-Y relationship in equation (1.2) is confounded by the unobserved ϵ_{1i} in (1.1) since it is an inseparable part of M.

As modeling techniques progress, standard mediation analysis is naturally absorbed into the linear structural equation models (SEM) (Iacobucci 2008). In recent years there is even a surge of research reconceptualizing direct and indirect effects from a counterfactual viewpoint (Imai, Keele, and Yamatomo 2010; Imai and Yamamoto 2013; Imai et al. 2011; Pearl 2009; 2012; Wang and Sobel 2013). Despite this most recent elaboration, the counterfactual causal mediation analysis still couches on the stringent assumption of uncorrelated error terms across equations.

II. Nonlinear Probability Models

When either or both of the mediator and outcome variables are discrete, nonlinear probability models such as logit/probit or generalized linear models (GLM) are more appropriate (McCullagh and Nelder 1989). Although these models have become indispensable parts of quantitative social science textbooks, endogenous explanatory variable in these nonlinear models poses severe challenges over and beyond their linear counterparts. Not only the familiar two-stage predictor substitution IV approach in linear models breaks down (Terza, Basu, and Rathouz

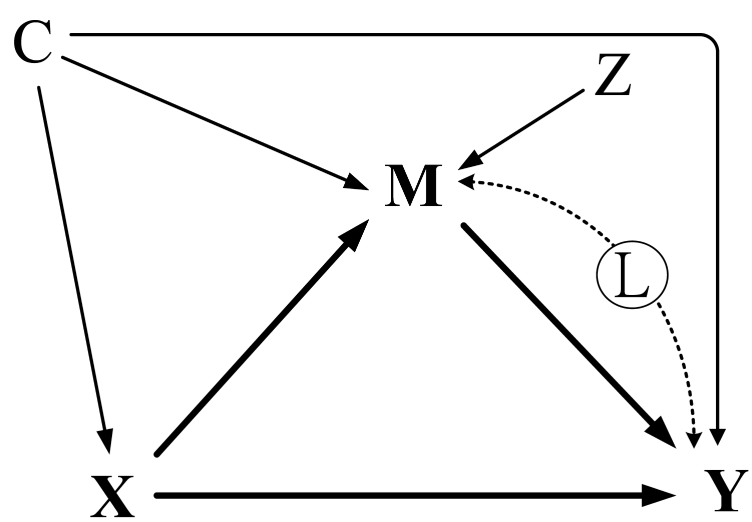
2008) but also the “product of coefficients” method in linear mediation analysis fails due to the often overlooked coefficient-rescaling problem in nonlinear probability models (Breen and Karlson 2013; Breen, Karlson, and Holm 2013; Karlson, Holm, and Breen 2012; Pearl 2012).

1. Breakdown of 2SLS Approach

The gut instinct to the endogenous regressor in nonlinear probability models such as logit or probit is indeed to follow the 2SLS template in linear models. That is, substitute the endogenous explanatory variable with its predictor constructed in the first stage regression and then run the appropriate nonlinear model in the second stage. For example, Lewis-Beck, Nadeau, and Elias (2008) used the so-called two-step probit by replacing the suspect endogenous regressor in probit model with its predicted value in the first-stage regression. Unfortunately, this intuitive “solution” to endogeneity is wrong. Terza, Basu, and Rathouz (2008) have proved the general inconsistency of 2SPS in the context of nonlinear models. Instead, they prove that an alternative two-stage IV approach in nonlinear models, called two-stage residual inclusion (2SRI), can obtain consistent estimation. The first-stage of this 2SRI estimator is identical to that of 2SPS, but the second-stage differs. Instead of substituting the endogenous regressor with its predictor, its actual observed value is maintained in the second-stage model while the residuals from the first-stage auxiliary regression are substituted for the unobserved or latent confounders as an additional “control variable,” hence the name 2SRI. Terza and his colleagues also prove that in linear models 2SRI=2SLS. Actually, some methods developed earlier for specific nonlinear models can be considered special cases of consistent 2SRI method. For example, the so-called control function approach proposed by Rivers and Vuong (1988) and generalized by Train (2009) and Wooldridge (2010) for continuous endogenous regressor in binary logit model is well known and widely used in applied research (e.g. Huang, Wang, and Lin 2012; 2013; Pertin and Train 2010).

If we classify the measurement of variables into five basic types: binary, nominal, ordinal, count, and continuous, then five types of endogenous regressor combine with five types of outcome variable result in 5x5=25 possible combinations. Up to this writing, only a small number of the 25 cells in this cross-table have methods proposed to handle endogeneity. Among them, the cell of continuous-continuous combination in linear models is undoubtedly the most researched and well developed. Consistency of 2SRI method proved by Terza and his colleagues is useful only when the residual in the first-stage regression is well defined. Unlike continuous dependent variables in linear models where residuals are simply the difference between observed and the predicted values, residuals of some categorical/discrete variables are not uniquely defined, with the exception of count variable (Hosmer and Lemeshow 2000; McCullagh and Nelder 1989). Therefore there is no single solution available for various types of nonlinear models. For example, ordinal variables can be coded arbitrarily so long as the coding preserves the order of categories. This makes the residuals vary according to different coding scheme.

Psychometricians Skrondal and Rabe-Hesketh (2004, 107-108) developed a "shared latent factor" approach closely related to Heckman's (1978) approach to endogenous binary regressor in the treatment effects model with continuous outcome. This shared factor approach introduces a latent factor, denoted as L in Figure 3, common to the error terms of the first- and second-stage regressions in order to induce their dependence. For more robust identification of parameters in the outcome equation of Y, an excluded instrument Z is often recommended in the first-stage equation.[2] This latent-factor approach can be a useful alternative to 2SRI when the residuals of discrete dependent variables, such as ordinal variables, are difficult to define. I will adopt this approach in the next section with an ordinal endogenous regressor in a binary outcome probit model, a combination received only scant attention in the existing literature (see Greene and Hensher 2010).



Source: By the Author.

Figure 3 A Mediation Model with Mediator-Outcome Confounding

[2] Unlike the linear case, here identification is theoretically achieved by the nonlinear functional forms. For practical purposes, however, an excluded instrument is recommended though not required.

2. Breakdown of Coefficient-Based Methods in Nonlinear Mediation Analysis

As discussed in the last section, in linear models mediation effect is estimated by comparing or multiplying coefficients estimated in two equations. In nonlinear models such as binary logit or probit and ordered logit or probit, comparing coefficients across equations is much more difficult. The challenges arise from the lack of separate identification of the mean and variance in these models. Although already shown in earlier literature (Maddala 1983; Winship and Mare 1984), this fact has often been overlooked until recent years when mediation and causal analyses are extended to nonlinear models (Breen and Karlson 2013; Breen, Karlson, and Holm 2013; Karlson, Holm, and Breen 2012). That is, the coefficient estimates of nonlinear probability models are equal to the underlying true coefficient, say β, divided by the unknown scale factor σ. True coefficients are identified only up to scale unless σ =1, which of course we do not know. Worse, the value of this scale factor not only varies in different equations with different outcome variables, but also changes in the same equation when one or more explanatory variables are added or excluded. This “coefficient-rescaling” feature explains why we *cannot* compare the coefficient of X from a logit or probit model excluding the mediator M with the corresponding coefficient from a logit or probit model including M. Likewise, multiplying coefficients from two different nonlinear equations makes little sense and cannot be interpreted as mediation effect of X on Y transmitted through M. Although coefficients’ signs and significance tests remain meaningful, they should not be used naively to estimate direct and indirect effects (Breen and Karlson 2013). Therefore, both the “product of coefficients” and “difference in coefficients” methods of estimating mediation effect in linear models collapse in the nonlinear probability models.

Many studies have devoted to the rescaling problem in nonlinear models. A group of scholars, mainly sociologists, try various ways to “standardize” coefficients in nonlinear models in order to parallel the two coefficient-based methods in linear models (Breen and Karlson 2013; Breen, Karlson, and Holm 2013; Karlson, Holm, and Breen 2012; Menard 2010; 2011). Despite tremendous efforts devoted to this end, all the methods proposed so far are limited only to the simplest models (such as mediation model with binary M and Y in Figure 1) and there has no consensus yet which rescaling method is the best.

Other researchers, however, choose to focus on the effects on probabilities. As the name "probability models" imply, the quantities of interest are the predicted probabilities of categories of the left-hand-side (LHS) variables. Furthermore, predicted probabilities are *not* sensitive to coefficient rescaling problem and easy to interpret (Imai, Keele, and Yamatomo 2010; Imai et al. 2011; Kuha and Goldthorpe 2010; Pearl 2012; Wooldridge 2010). For this school of scholars, consistent estimation of coefficients is of course important but only for the purpose of calculating predicted probabilities of the categorical outcome variables. The effect of a continuous explanatory variable on predicted probability can be evaluated by the average of marginal effects (Wooldridge 2010). For a categorical explanatory variable, its one-unit change effect (relative to its base category) on outcome is simply the average of discrete changes in predicted probability. Since effects on probability become the quantities of interest, researchers need to take an additional step to estimate the asymptotic standard errors by the delta-method (Greene 2012, 1123-1124) for significance tests of such effects. I adopt this probability-metric approach because it is more widely applicable to realistic nonlinear models and is easier to interpret.

III. Generalized SEM and an Empirical Illustration

The following sections intend to propose a strategy to integrate both sources of endogeneity, latent confounders and mediation, into a generalized structural equation model for discrete variables. I illustrate this modeling strategy with a simplified example of economic perceptions and vote choice using TEDS2012 postelection face-to-face interview survey data.[3] Our goal is not to settle the debates in this vast and rapidly growing literature but only to illustrate how best to build a more encompassing model so as to test competing theories while taking account both types of endogeneity.

1. Generalized SEM Approach

This section proposes an integrated parametric approach to the thorny problems of

[3] Data analyzed in this paper were from Taiwan's Election and Democratization Studies, 2012: Presidential and Legislative Elections (TEDS2012) (MOST 100-2420-H-002-030). The coordinator of multi-year project TEDS is Professor Chi Huang (National Chengchi University). TEDS2012 is a yearly project on the presidential and legislative elections in 2012. The principal investigator is Professor Yun-han Chu. More information is on TEDS website (http://www.tedsnet.org). The author appreciates the assistance in providing data by the institute and individuals aforementioned. The author alone is responsible for views expressed herein.

endogenous regressor in nonlinear probability models for cross-sectional data. This approach generalizes structural equation models by incorporating mediation analysis to deal with discrete mediators, on the one hand, and allowing correlated error terms to tackle unobserved confounders, on the other hand. The generalized SEM model-building consists of the following steps.

(1) Set up one separate equation for each potential discrete mediator variables M_j, $j=1, \ldots ,J$, in addition to the outcome equation for Y;
(2) Specify an appropriate probability threshold model for each equation with categorical left-hand-side (LHS) variable;
(3) Link, with a shared latent variable L, the error terms of the Y equation and the M_j equation that is most susceptible to unobserved confounders;
(4) Derive the covariance matrix of the error terms implied by the specified system of (J+1) equations so as to test for endogeneity due to unobserved variables;
(5) Estimate the system of (J+1) equations with maximum likelihood (ML) method; and
(6) Evaluate the effects of RHS variables on the predicted probabilities of the LHS variables.

Obviously, the first two steps are related to “mediation analysis” while the third and fourth steps are related to “shared latent-factor” approach. This generalized SEM approach is flexible enough to accommodate endogeneity for one-shot relationship in a single election. Admittedly, this generalized SEM approach has two limitations up to this writing. First, the third step of specifying shared latent factor is limited to one pair of error terms across two equations due to identification. Second, it is not applicable to dynamic data with a number of successive elections. Further generalizations allowing multiple pairs of latent factors and dynamic modeling remain to be developed in future studies.

2. Debates on the Relationship between Partisanship and Economic Perceptions

Economic voting model has established itself as a paradigm of studying electoral accountability based on past economic performance and future prospect (Kanji and Tannahill 2013; Lewis-Beck 1988; Lewis-Beck, Nadeau, and Elias 2008; Lewis-Beck and Stegmaier 2007; Lewis-Beck and Whitten 2013). The straightforward reward-punishment argument plus the valence of economic prosperity indeed make economy a key variable in many voting behavior research. However, factual economic condition may be a valence issue, subjective evaluation of economy may still be positional. Indeed recent "revisionist view" argues that economic voting is "endogenous" in the sense that partisanship strongly affects, if not distorts, voters' perceptions of macroeconomic performance (Anderson, Mendes, and Tverdova 2004; Duch 2008; Evans and Andersen 2006; Evans and Pickup 2010; Gerber and Huber 2010; Popescu 2013; Wlezien, Franklin, and Twiggs 1997). For example, incumbent party identifiers tend to evaluate the same objective economic conditions more favorably than opposition party identifiers. Economic perceptions may be largely partisan rationalization from this viewpoint. Kayser and Wlezien (2011) find that in Western Europe, economic conditions matter less when partisans proliferate and vice versa. In a recent special issue of *Electoral Studies* on economic voting, on the other hand, Lewis-Beck and Whitten (2013, 395) reassert that "...the economy places itself near the tip of the causal funnel."

Instead of taking sides a priori, we can cast the debates as competing models of the theoretical status of partisanship and economic perceptions. If the original economic voting paradigm is correct, then both partisanship and economic perceptions have only direct effects on voting choice. If the revisionist view is correct, than we need a separate equation to take account the mediation (or indirect) effect from partisanship to each economic perceptions and then from perceptions to voting decision. If neither of the separate equation finds significant effect of partisanship on economic perceptions, the revisionist view lacks empirical support and loses ground. But even if the effects of partisanship on economic perceptions are significant, the economic voting paradigm is not disconfirmed but only elaborated as a mediation mechanism. In any case, the key point here is that empirical test of the existence of mediation effect cannot be done in a single-equation format "followed by almost all practitioners in this field" (Lewis-Beck and Stegmaier 2007, 532). This comment also applies to economic voting research on Taiwan (see, for example, Ho et al., 2013; Hsieh, Lacy, and Niou 1998; Sheng 2009; Wang 2004; Wu and Lin 2012). Structural equation modeling is obviously a better alternative, although it has to be *generalized* to accommodate both categorical mediators and outcomes.

An ideal way out of this endogeneity-exogeneity controversy is to use panel data with all the explanatory variables measured at time earlier than the outcome variable. However, panel data are relatively rare. Furthermore, not all panel data are equally applicable. As Gerber and Huber (2010) point out, the time difference between two waves of survey interview should be short enough to ensure no intervening events occur to "contaminate" the pre-election measurements. This type of pre- and post-election panel data conducted in a short time interval is even rarer. Postelection cross-sectional surveys are most available, but the measurement of retrospective economic assessment after election is often questioned as contaminated by respondent's knowledge of who wins the election as well as the respondents' voting choice (Wu and Lin 2012). This may be related to the recall type of questions for both voting behavior and retrospective perception in postelection surveys. For example, respondents with poorer memory to recall the state of economy over the year before election may respond by "projecting" the image of the candidate s/he voted for backward in time. Furthermore, respondents voted for the elected candidate may have incentive to justify their choice by responding more positively to a retrospective economic evaluation question. Respondents' inability or unwillingness to recall correctly in postelection surveys can lead to correlation between their reported retrospective economic perception and voting choice. Unfortunately, neither ability nor willingness is observed by researchers. Here again, the confounding by unobservables cannot be relieved by simply controlling covariates in a single-equation format. We need to take account such possible endogeneity by allowing error terms across two involved equations to be correlated. The 2SRI and shared factor approach discussed in the last section are two candidates in the context of nonlinear outcome equation. I adopt the shared latent factor approach as discussed earlier. Again, this can be accomplished only by building a structural equation model.

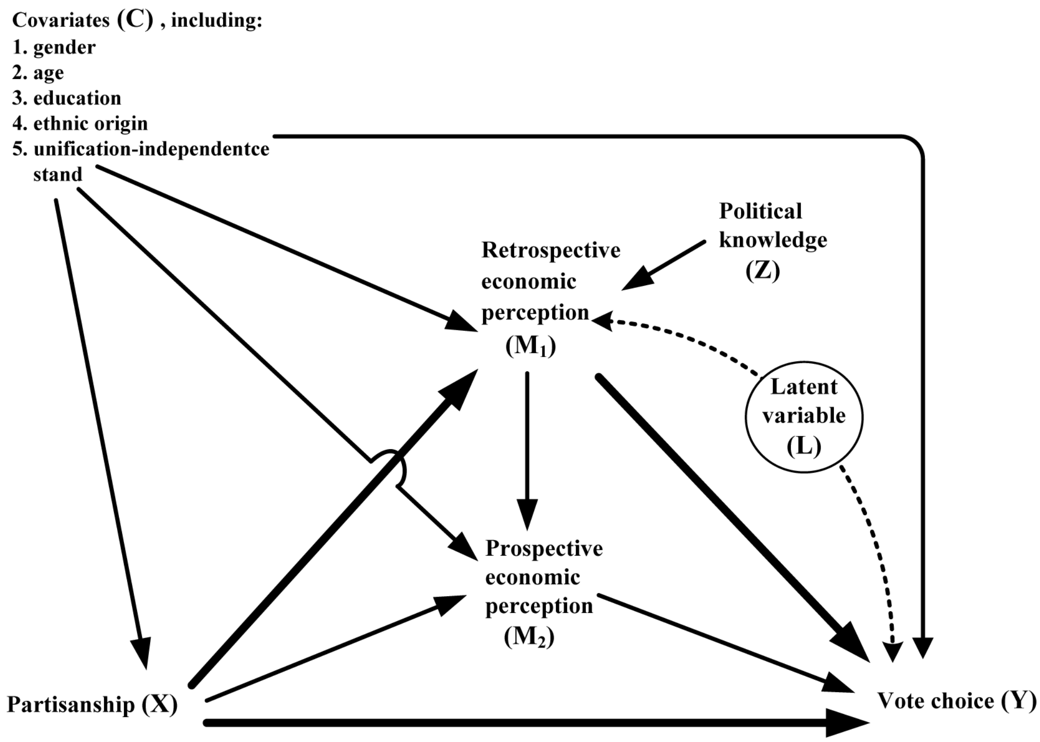
3. A Generalized Structural Equation Model with Two Related Mediators

To illustrate the suggested generalized SEM approach to tackle endogeneity in nonlinear models, I specify a simplified three-equation model to examine economic voting in Taiwan's 2012 presidential election. In this election, the incumbent KMT President Ma Ying-jeou was running for reelection against the opposition DPP candidate Tsai Ing-wen. The context of an incumbent president versus opposition challenger makes 2012 election suitable for testing the reward-punishment hypothesis of the economic voting paradigm.[4]

In order to test the revisionist view discussed earlier, I specify the sociotropic retrospective (M_1) and prospective (M_2) economic perceptions as possible mediators affected by partisanship (X), coded as pan-Blue, pan-Green, and independents. Following Lockerbie's (2008) argument, I allow the retrospective assessment affect the prospective perception. The outcome variable (Y) of course is the voting choice with incumbent Ma coded as 1. Our model specification can be

4 Only 38 respondents in the TEDS2012 sample reported voting for the PFP presidential candidate James Soong, so these cases are not included in our analysis.

illustrated as the diagram in Figure 4.



Source: By the Author.

Figure 4 A Generalized SEM of Partisanship, Economic Perceptions, and Voting

However, all the three left-hand-side (LHS) variables are categorical variables, with both economic perceptions coded as 3-category (worse, the same, better) ordered variables while the outcome variable, binary (see Appendix A for the coding of all the included variables). Appealing to the latent-variable framework, I assume the underlying propensity of choosing a particular category is continuous but unobserved. Our model can be formulated as a system of equations for three latent LHS variables with superscripts "*":

$$(2)\begin{cases} \text{retrospective:} & M_{1i}^{*} = \gamma_1 X_i + \gamma_2 Z_i + \boldsymbol{C}_i \boldsymbol{\gamma}_3 + u_{1i} \\ \text{prospective:} & M_{2i}^{*} = \beta_1 M_{1i}^{*} + \gamma_4 X_i + \boldsymbol{C}_i \boldsymbol{\gamma}_5 + u_{2i} \\ \text{vote choice:} & Y_i^{*} = \alpha + \beta_2 M_{1i}^{*} + \beta_3 M_{2i}^{*} + \gamma_6 X_i + \boldsymbol{C}_i \boldsymbol{\gamma}_7 + u_{3i} \end{cases}$$

where the vector $\boldsymbol{C}$ represents a vector of controlled covariates including social demographical variables of gender, age, education, and attitude toward the fundamental cleavage issue of unification with China vs. Taiwan independence. The identifying instrumental variable Z in the first equation is the degree of political knowledge, coded as low, middle, and high.

The relationship between each observed categorical and latent continuous variable is the familiar threshold model:

$$M_1 = \begin{cases} 1 & if \quad -\infty \le M_{1i}^* < \tau_{1,1} \\ 2 & if \quad \tau_{1,1} \le M_{1i}^* < \tau_{1,2} \\ 3 & if \quad \tau_{1,2} \le M_{1i}^* < \infty \end{cases}$$

$$M_2 = \begin{cases} 1 & if \quad -\infty \le M_{2i}^* < \tau_{2,1} \\ 2 & if \quad \tau_{2,1} \le M_{2i}^* < \tau_{2,2} \\ 3 & if \quad \tau_{2,2} \le M_{2i}^* < \infty \end{cases}$$

$$Y = \begin{cases} 1 & if \quad Y_i^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

In combination with the cumulative standard Normal distribution, the first two equations become familiar ordered probit, and the last is the binary probit. To build the model with possible M_1-Y confounding, I introduce a shared latent variable, L_i, to induce the dependence between the error terms of the first and the third equations, u_{1i} and u_{3i},

$$\begin{aligned} u_{1i} &= L_i + \varepsilon_{1i} \\ u_{3i} &= \lambda L_i + \varepsilon_{3i}, \end{aligned} \quad \text{where} \quad \begin{aligned} &L_i \sim N(0,1) \\ &\varepsilon_{1i} \sim N(0,1) \\ &\varepsilon_{3i} \sim N(0,1) \end{aligned}$$

Here L_i, ε_{1i} , and ε_{3i} are each independently Normally distributed with mean 0 and variance 1. For the sake of identification, I constrain the coefficient of L_i in the u_{1i} equation to be 1, but λ in the u_{3i} equation is a parameter to be estimated. Based on this specification, the covariance matrix of the error terms in these three equations is given by

$$\text{Cov}[(u_{1i}, u_{2i}, u_{3i})'] = \begin{bmatrix} 2 & 0 & \lambda \\ 0 & 1 & 0 \\ \lambda & 0 & \lambda^2 + 1 \end{bmatrix}$$

which implies that the correlation between u_{1i} and u_{3i} is

$$\rho = \frac{\lambda}{\sqrt{2(\lambda^2+1)}}.$$

Thus a test of H_0: ρ=0, or equivalently H_0: λ=0, is a test of endogeneity due to unobservables.

Suppose we ignore the possibility of both endogeneity problems, then the three-equation model above simplifies to the conventional single-equation binary probit model of vote choice. The results of this probit model, as shown in Appendix B, indicate that after controlling for all covariates there are no significant effects on voting choice of retrospective economic perception in terms of coefficients and probabilities. Only prospective perception of the state of the economy in the forthcoming year has positive and significant effect on voting for incumbent Ma. It is interesting to compare these results with the more encompassing three-equation model that allows for testing endogeneity.

4. Findings of Generalized SEM

Maximum likelihood (ML) estimates of our three-equation model (2)[5] are listed in Tables 1, 2, and 3 respectively. For brevity, I focus on results related to two possible sources of endogeneity. First of all, I examine whether economic perceptions are exogenous variables or endogenous mediators. As indicated in Tables 1 and 2, both retrospective and prospective economic assessments are strongly affected by partisanship, with incumbent blue-camp identifiers tend to evaluate the past and future economic conditions more favorably than opposition green-camp identifiers. A further joint test of significance of partisanship variable across the three equations obtains G^2=187.11 (p<0.001). This means that in Taiwan the partisan divide is so deep and wide that it shapes both citizens' evaluations of economy as well as presidential voting choice. Since this partisan divide has existed for more than two decades and constituted a long-term factor of citizens' political attitudes and behavior, I conclude that in Taiwan if economic perceptions affect voting choice at all they tend to mediate the strong effects of partisanship. Next, I examine the existence of confounding factor. As shown at the bottom of Table 3, $\hat{\lambda}$=-2.147 and the test of H_0: λ=0 is rejected at conventional level (p<0.05). In terms of correlation between u_{1i} and u_{3i}, $\hat{\rho}$=-.641 (p<0.001) which also indicates that the residuals of the retrospective economic perception and voting choice are correlated. These results confirm the necessity of incorporating mediation analysis as well as the latent confounding factor into our model.

5 The model was estimated with Stata 13.1. Stata commands are listed in Appendix C.

Table 1 Determinants of Retrospective Economic Perception

	Coefficient estimates	**S.E.**	***p*-value**
Political knowledge(middle=0)			
low	0.088	(0.120)	0.465
high	-0.132	(0.122)	0.282
Partisanship(non-partisan=0)			
blue	0.742***	(0.143)	<0.001
green	-0.858***	(0.157)	<0.001
Gender(female=0)			
male	0.123	(0.102)	0.225
Age	-0.004	(0.004)	0.285
Education(low=0)			
middle	0.279†	(0.149)	0.061
high	0.471**	(0.156)	0.003
Origin(Hakka=0)			
Taiwanese	-0.044	(0.143)	0.755
Mainlander	0.286	(0.184)	0.120
Unification-Independence(status quo=0)			
unification	-0.041	(0.144)	0.775
independence	-0.048	(0.133)	0.716
Latent Variable	1(constrained)		
cutpoint 1	-0.149	0.314	0.636
cutpoint 2	1.576***	0.317	<0.001

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

Table 2 Determinants of Prospective Economic Perception

	Coefficient estimates	**S.E.**	***p*-value**
Retrospective(same=0)			
worse	-0.788***	(0.083)	<0.001
better	0.393***	(0.096)	<0.001
Partisanship(non-partisan=0)			
blue	0.298**	(0.103)	0.004
green	-0.227*	(0.111)	0.040
Gender(female=0)			
male	-0.033	(0.070)	0.638
Age	-0.001	(0.003)	0.751
Education(low=0)			
middle	-0.024	(0.104)	0.815
high	0.032	(0.105)	0.760
Origin(Hakka=0)			
Taiwanese	0.048	(0.102)	0.640
Mainlander	0.170	(0.133)	0.203
Unification-Independence(status quo=0)			
unification	0.184†	(0.103)	0.076
independence	-0.241**	(0.092)	0.009
cutpoint 1	-0.863***	(0.219)	<0.001
cutpoint 2	0.701**	(0.219)	0.001

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

Table 3 Determinants of Presidential Voting Choice

	Coefficient estimates	S.E.	*p*-value
Retrospective(same=0)			
worse	-2.328†	(1.199)	0.052
better	2.044†	(1.122)	0.068
Prospective(same=0)			
worse	-0.471	(0.374)	0.207
better	0.979†	(0.566)	0.084
Partisanship(non-partisan=0)			
blue	2.668**	(0.937)	0.004
green	-2.827**	(1.022)	0.006
Gender(female=0)			
male	-0.448	(0.335)	0.182
Age	0.006	(0.012)	0.626
Education(low=0)			
middle	-0.421	(0.460)	0.360
high	-0.382	(0.472)	0.418
Origin(Hakka=0)			
Taiwanese	-0.756	(0.492)	0.124
Mainlander	1.327	(0.897)	0.139
Unification-Independence(status quo=0)			
unification	0.424	(0.505)	0.401
independence	-0.757	(0.463)	0.102
Latent Variable ($\hat{\lambda}$)	-2.147*	(1.087)	0.048
constant	1.895	(1.189)	0.111
Number of observations=1,122			
Log-likelihood=-2288.9259			

Data: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

As explained in the last section, I use probability metric to evaluate and interpret the effects of explanatory variables in each equation on the LHS categorical variables. The two LHS variables in the first two equations are three-category (worse, the same, better) ordinal variables and rigorously speaking requires six tables with each contains "effects on probability" as well as standard errors estimated by delta-method for each category. For the sake of parsimony, however, I list only the effect on probability of choosing the highest category, i.e., perceiving the economy has gotten better and will become better in Table 4 and Table 5, respectively. The final outcome variable of vote choice is binary and thus requires only one table, as listed in Table 6. Both Tables 4 and 5 indicate that in Taiwan the revisionist view is correct since partisanship significantly affects both retrospective and prospective economic perceptions. For example, on average, Blue-camp supporters are 0.17 more likely than non-partisans to believe the economy has gotten better over the past year while the Green-camp supporters are 0.07 less likely to believe so. Furthermore, as shown in Table 5, besides the same direction of partisan bias toward the prospective economic perception, optimistic/pessimistic retrospective assessment also tends to transmit to the prospective perception. After taking account partisan effects on economic perceptions as well as the latent confounding factor, Table 6 indicates highly significant effects of retrospective perceptions on voting choice as expected by the punishment-reward hypothesis of economic voting. This finding is stark different from the conventional single-equation probit model in Appendix B. On average those who consider the economy has gotten better have 0.159 higher probability of voting for incumbent President Ma while those who consider economy has gotten worse have 0.142 lower probability of voting for Ma. Curiously enough, prospective perception has much weaker effect on voting choice and is limited to rewarding effect. Overall, partisanship is indeed the most important determinant of vote choice, with economic voting tracing not far behind. Social and political cleavages, such as unification with China vs. Taiwan independence, are still significant but with effect on probability somewhat smaller than retrospect voting.

Table 4 Effects of Explanatory Variables on the Probability of Perceiving The State of the Economy Has Gotten "Better" over the Past Year

	Effects on probability	**Delta-method S.E.**	***p*-value**
Political knowledge(middle=0)			
low	0.017	(0.023)	0.465
high	-0.025	(0.024)	0.282
Partisanship(non-partisan=0)			
blue	0.170***	(0.029)	<0.001
green	-0.070***	(0.019)	<0.001
Gender(female=0)			
male	0.024	(0.020)	0.225
Age	-0.001	(0.001)	0.285
Education(low=0)			
middle	0.047†	(0.024)	0.055
high	0.086**	(0.027)	0.001
Origin(Hakka=0)			
Taiwanese	-0.008	(0.027)	0.758
Mainlander	0.061	(0.039)	0.122
Unification-Independence(status quo=0)			
unification	-0.008	(0.027)	0.773
independence	-0.009	(0.025)	0.712

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

Table 5　Effects of Explanatory Variables on the Probability of Perceiving The State of the Economy Will Get “Better” in the Forthcoming Year

	Effects on probability	**Delta-method S.E.**	***p*-value**
Retrospective(same=0)			
worse	-0.183***	(0.020)	<0.001
better	0.135***	(0.034)	<0.001
Partisanship(non-partisan=0)			
blue	0.081**	(0.027)	0.003
green	-0.052*	(0.026)	0.045
Gender(female=0)			
male	-0.008	(0.018)	0.638
Age	-0.0002	(0.0007)	0.751
Education(low=0)			
middle	-0.006	(0.026)	0.815
high	0.008	(0.027)	0.759
Origin(Hakka=0)			
Taiwanese	0.012	(0.025)	0.636
Mainlander	0.044	(0.035)	0.205
Unification-Independence(status quo=0)			
unification	0.050†	(0.029)	0.085
independence	-0.059**	(0.022)	0.007

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

Table 6 Effects of Explanatory Variables on the Probability of Voting for Ma

	Effects on probability	**Delta-method S.E.**	***p*-value**
Retrospective(same=0)			
worse	-0.142***	(0.032)	<0.001
better	0.159**	(0.061)	0.009
Prospective(same=0)			
worse	-0.027	(0.019)	0.156
better	0.051*	(0.022)	0.022
Partisanship(non-partisan=0)			
blue	0.336***	(0.043)	<0.001
green	-0.358***	(0.053)	<0.001
Gender(female=0)			
male	-0.023	(0.015)	0.117
Age	0.0003	(0.0006)	0.618
Education(low=0)			
middle	-0.021	(0.021)	0.311
high	-0.019	(0.022)	0.372
Origin(Hakka=0)			
Taiwanese	-0.041†	(0.021)	0.056
Mainlander	0.075	(0.046)	0.103
Unification-Independence(status quo=0)			
unification	0.023	(0.026)	0.374
independence	-0.045*	(0.022)	0.042

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

IV. Conclusions

This study addresses the ubiquitous endogeneity problem in empirical social science research. Against the backdrop of recent development in quantitative methodology in different disciplines, I review some overlooked pitfalls while dealing with nonlinear probability models. I point out that nonlinearity is a generic term with diversified functional forms and that there is no single panacea yet for endogeneity problems in various types of nonlinear models. Each possible combination of endogenous regressor and outcome variables deserves a careful examination on its own. It is extremely dangerous to simply transplant methods developed for linear models to nonlinear models.

That said, this study draws attention to the generalized SEM strategy which goes beyond the single-equation format. This approach can take account endogeneity caused by latent factors and mediation effect in the context of an ordinal endogenous regressor in binary outcome probit model. I hope that this new modeling strategy can ameliorate the endogeneity problem in certain frequently used nonlinear probability models and shed new light on the exogeneity-endogeneity debates in applied research.

* * *

Received: 2014.05.05; Revised: 2014.07.07; Accepted: 2014.10.02

Appendix A Coding of Variables

(Sociotropic Retrospective Economic Perception) Would you say that over the past year, the state of the economy of Taiwan has gotten better, stayed about the same, or gotten worse?

1. Worse
2. About the same
3. Better

(Sociotropic Prospective Economic Perception) Would you say that in the forthcoming year, the state of the economy of Taiwan will get better, stay about the same, or get worse?

1. Worse
2. About the same
3. Better

(Voting Choice) Which candidate did you vote for?

0. TSAI Ing-wen and SU Jia-chyuan
1. MA Ying-jeou and WU Den-yih

(Partisanship)

1. Pan-Blue (KMT+PFP+NP)
2. Pan-Green (DPP+TSU)
3. Independents (non-partisans)

(Unification-Independence) Concerning the relationship between Taiwan and mainland China, which of the following positions do you agree with:

1. Unification with China
2. Status quo
3. Taiwan independence

(Ethnic Origin) Father's ethnic background

1. Taiwanese Hakka
2. Taiwanese Min-Nan
3. Mainlander

(Education) Educational level

1. Primary school and below
2. High school
3. College and above

(Political Knowledge) Level of political knowledge

Based on the number of correct answers to the following six items:

- Who is the current president of the United States?
- Who is the current premier of our country?
- What institution has the power to interpret the constitution?
- What was the unemployment rate in Taiwan as of the end of last year?
- Which party came in second in seat in the Legislative Yuan?
- Who is the current Secretary-General of the United Nations – Kofi Annan, Kurt Waldheim, Ban Ki-moon, or Boutros Boutros-Ghali?

1. Low (0-2 items correct)
2. Middle (3-4 items correct)
3. High (5-6 items correct)

Appendix B Conventional Single-Equation Probit Model

	Coefficient estimates	S.E.	Effects on probability	Delta-method S.E.
Political knowledge (middle=0)				
low	0.247	(0.163)	0.026	(0.017)
high	0.027	(0.187)	0.003	(0.020)
Retrospective (same=0)				
worse	-0.122	(0.158)	-0.013	(0.018)
better	0.070	(0.206)	0.007	(0.022)
Prospective (same=0)				
worse	-0.221	(0.164)	-0.025	(0.019)
better	0.477*	(0.203)	0.055*	(0.024)
Partisanship (non-partisan=0)				
blue	1.752***	(0.172)	0.374***	(0.037)
green	-1.699***	(0.162)	-0.466***	(0.043)
Gender (female=0)				
male	-0.183	(0.137)	-0.020	(0.015)
Age	0.001	(0.006)	0.0001	(0.0006)
Education (low=0)				
middle	-0.083	(0.204)	-0.009	(0.022)
high	0.055	(0.218)	0.006	(0.023)
Origin (Hakka=0)				
Taiwanese	-0.430*	(0.190)	-0.049*	(0.022)
Mainlander	0.745†	(0.383)	0.091†	(0.049)
Unification-Independence (status quo=0)				
unification	0.183	(0.235)	0.021	(0.027)
independence	-0.409*	(0.160)	-0.047*	(0.019)
constant	0.515	(0.426)	--	--
Number of observations=1,122				
Log-likelihood=-219.67586				

Source: TEDS2012.

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; †: $p < 0.1$.

Appendix C Stata Program for Implementing Three-Equation Economic Voting Model

```
* complete observations
mark nomiss
markout nomiss pvote2 sretro spros bgpid3 sex agey edu3 sengi3 tondu3

* generalized SEM
gsem (sretro <- pk_l pk_h ib3.bgpid3 ib2.sex agey ib1.edu3 ib1.sengi3 ib2.tondu3 L@1, oprobit) ///
    (spros <- ib2.sretro ib3.bgpid3 ib2.sex agey ib1.edu3 ib1.sengi3 i.tondu3, oprobit) ///
    (pvote2 <- ib2.sretro ib2.spros ib3.bgpid3 ib2.sex agey ib1.edu3 ib1.sengi3 ib2.tondu3 L, probit) ///
    if nomiss==1, var(L@1)

* endogeneity test of H0: rho=0
nlcom (rho: _b[pvote2:L]/(sqrt(2)*sqrt(1+_b[pvote2:L]^2)))
```

References

Anderson, Christopher J., Silvia M. Mendes, and Yuliya V. Tverdova. 2004. "Endogenous Economic Voting: Evidence from the 1997 British Election." *Electoral Studies* 23(4): 683-708.

Baron, Reuben M., and David A. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51(6): 1173-1182.

Bollen, Kenneth A. 2012. "Instrumental Variables in Sociology and the Social Sciences." *Annual Review of Sociology* 38: 37-72.

Breen, Richard, and Kristian B. Karlson. 2013. "Counterfactual Causal Analysis and Nonlinear Probability Models." In *Handbook of Causal Analysis for Social Research*, ed. Stephen L. Morgan. Heidelberg: Springer.

Breen, Richard, Kristian B. Karlson, and Anders Holm. 2013. "Total, Direct, and Indirect Effects in Logit and Probit Models." *Sociological Methods & Research* 42(2): 164-191.

Duch, Raymond M., and Randolph T. Stevenson. 2008. *The Economic Vote: How Political and Economic Institutions Condition Election Results*. Cambridge: Cambridge University Press.

Evans, Geoffrey, and Robert Andersen. 2006. "The Political Conditioning of Economic Perceptions." *The Journal of Politics* 68(1):194-207.

Evans, Geoffrey, and Mark Pickup. 2010. "Reversing the Causal Arrow: The Political Conditioning of Economic Perceptions in the 2000–2004 U.S. Presidential Election Cycle." *The Journal of Politics* 72(4):1236-1251.

Gerber, Alan S., and Gregory A. Huber. 2010. "Partisanship, Political Control, and Economic Assessments." *American Journal of Political Science* 54(1): 153-173.

Greene, William H. 2012. *Econometric Analysis* (7th ed.). Essex: Pearson Education.

Greene, William H., and David A. Hensher. 2010. *Modeling Ordered Choices: A Primer*. New York: Cambridge University Press.

Hayes, Andrew F. 2013. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: The Guilford Press.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46(6): 931-959.

Ho, Karl, Harold D. Clarke, Li-khan Chen, and Dennis Lu-chung Weng. 2013. "Valence Politics and Electoral Choice in a New Democracy: The Case of Taiwan." *Electoral Studies* 32(3): 476-481.

Hosmer, David W., and Stanley Lemeshow. 2000. *Applied Logistic Regression* (2nd ed.). New York: Wiley.

Hsieh, John Fuh-sheng, Dean Lacy, and Emerson M. S. Niou. 1998. "Retrospective and Prospective Voting in a One-Party-Dominant Democracy: Taiwan's 1996 Presidential Election." *Public Choice* 97(3): 383-399.

Huang, Chi, Hung-chung Wang, and Chang-chih Lin. 2012. "Knowledge of the Electoral System and Voter Turnout." *Taiwanese Political Science Review* 16(1): 239-279.

------. 2013. "Knowledge of the Electoral System and Voting: Taiwan's 2008 and 2012 Legislative Elections." *Issues & Studies* 49(4): 1-45.

Iacobucci, Dawn. 2008. *Mediation Analysis*. Thousand Oaks, CA: Sage Publications.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4):765-789.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1): 51-71.

Imai, Kosuke, and Teppei Yamamoto. 2013. "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments." *Political Analysis* 21(2): 141-171.

Jackson, John E. 2008. "Endogeneity and Structural Equation Estimation in Political Science." In *The Oxford Handbook of Political Methodology*, eds. Janet Box-Steffensmeier, Henry E. Brady and David Collier. Oxford: Oxford University Press.

Kanji, Mebs, and Kerry Tannahill. 2013. "What Are the Current Preoccupations of Economic Voting Research?" *Electoral Studies* 32(3): 391-392.

Karlson, Kristian B., Anders Holm, and Richard Breen. 2012. "Comparing Regression Coefficients Between Same-Sample Nested Models Using Logit and Probit: A New Method." *Sociological Methodology* 42(1): 286-313.

Kayser, Mark A., and Christopher Wlezien. 2011. "Performance Pressure: Patterns of Partisanship and the Economic Vote." *European Journal of Political Research* 50(3):365-394.

Kuha, Jouni, and John H. Goldthorpe. 2010. "Path Analysis for Discrete Variables: The Role of Education in Social Mobility." *Journal of the Royal Statistical Society, Series A* 173(2):

351-369.

Lewis-Beck, Michael S. 1988. *Economics and Elections: The Major Western Democracies*. Ann Arbor: University of Michigan Press.

Lewis-Beck, Michael S., Richard Nadeau, and Angelo Elias. 2008. “Economics, Party, and the Vote: Causality Issues and Panel Data.” *American Journal of Political Science* 52(1): 84-95.

Lewis-Beck, Michael S., and Mary Stegmaier. 2007. “Economic Models of Voting.” In *The Oxford Handbook of Political Behavior*, eds. Russell J. Dalton and Hans-Dieter Klingemann. Oxford: Oxford University Press.

Lewis-Beck, Michael S., and Guy D. Whitten. 2013. “Economics and Elections: Effects Deep and Wide.” *Electoral Studies* 32(3): 393-395.

Lockerbie, Brad. 2008. *Do Voters Look to the Future? Economics and Elections*. Albany, NY: State University of New York Press.

MacKinnon, David P. 2008. *Introduction to Statistical Mediation Analysis*. New York: Lawrence Erlbaum Associates.

McCullagh, P., and John A. Nelder. 1989. *Generalized Linear Models* (2nd ed.). London: Chpaman and Hall.

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

Menard, Scott. 2010. *Logistic Regression: From Introductory to Advanced Concepts and Applications*. Thousand, Calif.: SAGE.

------. 2011. “Standards for Standardized Logistic Regression Coefficients.” *Social Forces* 89(4): 1409-1428.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge: Cambridge University Press.

------. 2012. “The Causal Foundations of Structural Equation Modeling.” In *Handbook of Structural Equation Modeling*, ed. Rick H. Hoyle. New York: The Guilford Press.

Pertin, Amil, and Kenneth Train. 2010. “A Control Function Approach to Endogeneity in Consumer Choice Models.” *Journal of Marketing Research* 47(1): 3-13.

Popescu, Gheorghe H. 2013. “Partisan Differences in Evaluations of the Economy.” *Economics, Management and Financial Markets* 8(1): 130-135.

Rivers, Douglas, and Quang H. Vuong. 1988. “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models.” *Journal of Econometrics* 39(3): 347-366.

Sheng, Sing-yuan. 2009. "The Impacts of Economic and Welfare Issues on Voting Behavior: A Discussion of 2008 Presidential Election." (in Chinese) In *2008 Taiwan Presidential Election: The Second Party Turnover*, eds. Lu-hui Chen, Qing-xin You and Chi Huang. Taipei: Wunan.

Skrondal, Anders, and Sophia Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, Fla.: Chapman & Hall/CRC.

Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1): 188-200.

Terza, Joseph V., Anirban Basu, and Paul J. Rathouz. 2008. "Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling." *Journal of Health Economics* 27(3): 531-543.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation* (2nd ed.). New York: Cambridge University Press.

Wang, Bo-yauo. 2004. "The Evaluation of Economic Voting and Voting Choice: 2001 Legislative Election of Taiwan." (in Chinese) *Journal of Electoral Studies* 11(1): 171-195.

Wang, Xialu, and Michael E. Sobel. 2013. "New Perspectives on Causal Mediation Analysis." In *Handbook of Causal Analysis for Social Research*, ed. Stephen L. Morgan. Dordrecht: Springer.

Winship, Christopher, and Robert D. Mare. 1984. "Regression Models with Ordinal Variables." *American Sociological Review* 49(4): 512-525.

Wlezien, Christopher, Mark Franklin, and Daniel Twiggs. 1997. "Economic Perceptions and Vote Choice: Disentangling the Endogeneity." *Political Behavior* 19(1): 7-17.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge: The MIT Press.

Wu, Chin-en, and Yi-tze Lin. 2012. "Economic Voting and Presidential Elections: An Assessment of Validity and Endogeneity." (in Chinese) *Taiwanese Political Science Review* 16(2): 175-232.

非線性機率模型中的內因自變數問題：廣義結構式模型及其應用

黃紀 *

《本文摘要》

社會科學研究中，解釋變數常發生棘手的內因 (endogeneity) 問題。線性模型之內因自變數處理方式，如工具變數及其延伸，討論頗多。但若依變數爲類別變數，其非線性的機率模型面對內因自變數，問題遠比線性模型複雜得多，絕不宜盲目以線性模型的處理方式比照適用。本文的目的，在釐清內因問題的起源，並區分實證研究較常遇到的兩大內因來源：未觀測到的潛在因素及內因中介變數，回顧線性模型文獻中對兩者的因應方式，並分析這些方法在非線性機率模型中面臨的困難與挑戰。接著本文提出廣義結構式模型的解決方案，既可同時因應兩種內因問題，亦能兼顧非線性模型的統計特性。爲了說明廣義結構式模型的應用，本文舉經濟投票文獻中對「整體經濟回顧與前瞻型評價」的內因性辯論爲例，建立能兼容相競學理的廣義結構式模型，並以實證資料 TEDS2012 進行檢驗，發現回顧型評價對投票抉擇有顯著的影響。

關鍵詞：內因性、非線性機率模型、中介變數分析、係數之跨模型比較、經濟投票

* 國立政治大學講座教授、政治學系教授暨選舉研究中心合聘研究員。