DETERMINING THE PROPER SPECIFICATION FOR ENDOGENOUS COVARIATES IN DISCRETE DATA SETTINGS

Angela Vossmeyer

Department of Economics, University of California, Irvine, CA, USA

ABSTRACT

An important but often overlooked obstacle in multivariate discrete data models is the specification of endogenous covariates. Endogeneity can be modeled as latent or observed, representing competing hypotheses about the outcomes being considered. However, little attention has been applied to deciphering which specification is best supported by the data. This paper highlights the use of existing Bayesian model comparison techniques to investigate the proper specification for endogenous covariates and to understand the nature of endogeneity. Consideration of both observed and latent modeling approaches is emphasized in two empirical applications. The first application examines linkages for banking contagion and the second application evaluates the impact of education on socioeconomic outcomes.

Bayesian Model Comparison
Advances in Econometrics, Volume 34, 223–247

ISSN: 0731-9053/doi:10.1108/S0731-905320140000034010

Keywords: Bayesian estimation; data augmentation; educational attainment; financial contagion; latent data; marginal likelihood

JEL classifications: C35; C52; G01; I00

1. INTRODUCTION

In empirical applications, endogenous regressors are generally the key variables of interest. Treatment models, triangular systems with recursive endogeneity, and sequential decision-making all feature endogenous covariates that often represent the main components of the study. In continuous data settings, modeling endogeneity is simple and interpretation is straightforward. In discrete data settings, modeling endogeneity is complicated because it can take several forms based on latent or observed data. This is not a limitation of the system. Instead, this feature increases the flexibility of such models because unobservables can be captured as explanatory variables separately from variables that are observable to the econometrician. This is only beneficial if a formal model comparison can be performed to decipher which depicted relationship is best supported by the data and to resolve competing hypotheses about the type of endogeneity. Investigating both approaches strengthens the eventual results by increasing a researcher's understanding of the relationships and the dependence structure being modeled. Model testing is easily afforded by existing Bayesian model comparison techniques.

Initial latent data modeling innovations occurred in psychometrics, where the traditional usage of latent variables focused on measurement error and hypothetical constructs (Muthen, 2002). In econometrics, latent data analysis advances discrete choice methods in which choice outcomes are linked to latent utility. For a review, see Jeliazkov and Rahman (2012). Bayesian econometrics further benefited by the association between data augmentation and latent variables (Tanner & Wong, 1987). Although latent modeling approaches have captured a variety of statistical concepts, including random coefficients, missing data, discrete choice, and finite mixture modeling, latent variables are rarely employed as regressors since an investigator has not or can not measure or observe them. Macroeconomics has moved in this direction with factor models; however, latent covariates remain unexplored in applied microeconomic research. Furthermore, across all fields, little attention has been applied to formally compare such models. In a recent marketing paper, Mintz, Currim, and Jeliazkov (2013) look at both specifications and find that a latent measure of information processing pattern better explains an individual's propensity to buy. Overlooking the consideration of both approaches can lead to specification errors and misrepresent the relationships being examined.

This paper employs Bayesian model selection methods for comparing latent and observed endogeneity models in two empirical applications. Each application features competing hypotheses discussed in the literature and a formal motivation for using observed or latent endogeneity. The first application examines banking contagion and the relative influence or spread of contagion from both regional and network linkages. The second application considers the impact of education on adult socioeconomic status. These applications highlight a key aspect of this research topic. While an applied researcher may have *a priori* expectations of the "correct" model, in most cases and especially in these examples, arguments for both approaches are easily formed, making it difficult for a researcher to completely rule out a specification without performing model comparison.

The rest of this paper is organized as follows: Section 2.1 reviews each specification in a simple bivariate system of equations and Section 2.2 discusses existing techniques for model selection. Section 3 considers the application to financial contagion and Section 4 considers the application to education. Finally, Section 5 offers concluding remarks.

2. METHODOLOGY

2.1. Model

To exemplify the approaches discussed in this paper, consider a bivariate model with recursive endogeneity, where latent data are referred to as $\mathbf{y}_i^*$ and observed data are referred to as $\mathbf{y}_i$. The two different modeling techniques, latent and observed, are shown as:

Observed Endogeneity

$$\begin{aligned} y_{i1}^* &= \mathbf{x}'_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1} \\ y_{i2}^* &= \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + y_{i1}\gamma_2 + \varepsilon_{i2} \end{aligned} \tag{1}$$

Latent Endogeneity

$$\begin{aligned} y_{i1}^* &= \mathbf{x}'_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1} \\ y_{i2}^* &= \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + y_{i1}^*\gamma_2 + \varepsilon_{i2}. \end{aligned} \tag{2}$$

The latent data are related to the observed outcomes by a link function depending on the values y_{ik} can take, for equations $k = 1, 2$. The binary setting occurs when $y_{ik} = 1\{y_{ik}^* > 0\}$, and the censored setting occurs when $y_{ik} = y_{ik}^* \cdot 1\{y_{ik}^* > 0\}$. The link function for ordered data is $y_{ik} = \sum_{j=1}^{J} 1\{y_{ik}^* > \alpha_{k,j-1}\}$ for J ordered alternatives, where α_{kj} is a cut-point between the categories. In this context, the observed endogeneity system differs from (1) and instead is:

$$\begin{aligned} y_{i1}^* &= \mathbf{x}'_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1} \\ y_{i2}^* &= \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + 1\{y_{i1} = 2\}\gamma_{22} + 1\{y_{i1} = 3\}\gamma_{23} + \cdots + 1\{y_{i1} = J\}\gamma_{2J} + \varepsilon_{i2}, \end{aligned} \tag{3}$$

where there is a set of endogenous indicator variables for $J-1$ categories, as opposed to a single endogenous regressor as in (1). This case is explored in the second application to education in Section 4. For simplicity, assume $\varepsilon_i \equiv (\varepsilon_{i1}, \varepsilon_{i2})' \sim N_2(0, \Omega)$ and $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$ in Eqs. (1), (2), and (3). Models with endogeneity have been difficult to estimate when the response variables of interest are not continuous because standard two-stage estimators are inapplicable in this context. Therefore, estimation in this paper relies on Markov chain Monte Carlo (MCMC) methods, which are discussed in detail in each application.

The latent data have the customary random utility interpretation underlying the theory on discrete choice analysis in econometrics. Therefore, even though the observed data can only take certain values, the latent variables that determine those outcomes are unrestricted. The fact that the latent utilities can be changed without necessarily inducing a corresponding change in the observed variable is a key distinction between these models (Mintz et al., 2013). Furthermore, the observed and latent specifications pose different relationships between the variables of interest because latent data measure intentions and observed data measure actual actions or outcomes. In (2), latent endogeneity says that intentions about y_{i1} determine intentions about y_{i2}. In (1), observed endogeneity says that actions about y_{i1} determine intentions about y_{i2} (Maddala, 1983). Despite the clear interpretation of each model, in most cases, it is difficult for a researcher to decipher which specification is correct. Generally, convincing hypotheses or arguments can be made in support of either modeling approach, hence motivating the need for model comparison.

It is important to note that these considerations extend to larger systems of equations, models for sample selection, potential outcomes, simultaneous equations, and more. For a review of some of these models, see van Hasselt (2014) and Li and Tobias (2014). Any multivariate discrete outcome model should not overlook this problem. The bivariate system is considered here to stress the importance of the issue, highlight the ease of considering both models, and offer a more complete understanding of the relationships in each application.

2.2. Model Comparison

Model comparison techniques are often employed to deal with issues of model uncertainty and variable selection. This paper utilizes these same approaches to determine the nature of endogeneity. The methods used in this paper are from Chib (1995) and Chib and Jeliazkov (2001), which are computationally convenient and do not require much additional coding. The applications in this paper span a number of discrete choice models, including ordered probit, Tobit, and binary probit. Therefore, the model comparison methods discussed here are general across these classes of models.

Given the data $\mathbf{y}$, interest centers upon the models $\{\mathcal{M}_l, \mathcal{M}_o\}$ where $\mathcal{M}_l$ represents the latent endogeneity model and $\mathcal{M}_o$ represents the observed endogeneity model. Each model is characterized by a sampling density $\{f(\mathbf{y}|\mathcal{M}_l, \boldsymbol{\theta}_l), f(\mathbf{y}|\mathcal{M}_o, \boldsymbol{\theta}_o)\}$ where $\{\boldsymbol{\theta}_l, \boldsymbol{\theta}_o\}$ are model-specific parameter vectors. Bayesian model selection proceeds by comparing the models through their posterior odds ratio

$$\frac{\Pr(\mathcal{M}_l|\mathbf{y})}{\Pr(\mathcal{M}_o|\mathbf{y})} = \frac{\Pr(\mathcal{M}_l)}{\Pr(\mathcal{M}_o)} \times \frac{m(\mathbf{y}|\mathcal{M}_l)}{m(\mathbf{y}|\mathcal{M}_o)}. \tag{4}$$

Chib (1995) recognized the *basic marginal likelihood identity* in which the marginal likelihood for model $\mathcal{M}_l$ can be expressed as

$$m(\mathbf{y}|\mathcal{M}_l) = \frac{f(\mathbf{y}|\mathcal{M}_l, \boldsymbol{\theta}_l)\pi(\boldsymbol{\theta}_l|\mathcal{M}_l)}{\pi(\boldsymbol{\theta}_l|\mathbf{y}, \mathcal{M}_l)}. \tag{5}$$

Calculation of the marginal likelihood is then reduced to finding an estimate of the posterior ordinate $\pi(\boldsymbol{\theta}_l^*|\mathbf{y}, \mathcal{M}_l)$ at a single point $\boldsymbol{\theta}_l^*$, which is often taken as the posterior mean or mode. Since the topics in this paper involve multivariate discrete data, sampling densities are often analytically intractable. A straightforward approach for evaluating the likelihood function employed in this paper is the Chib–Ritter–Tanner (CRT) method, which was developed in Jeliazkov and Lee (2010).

Decomposition of the posterior ordinate varies across the examples in this paper, with the most difficult being the multivariate ordered probit model used in the education application. This case is a bit more complex due to the additional cut-point parameters $\boldsymbol{\delta}$.[1] This section outlines the decomposition used for the ordered probit model, which follows from Jeliazkov, Graves, and Kutzbach (2008). It should be noted that the decomposition for the binary probit and Tobit models are simplified versions of the ordered probit without computations for the $\boldsymbol{\delta}$ parameter vector. Let $\boldsymbol{\lambda} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ be a parameter vector for all endogenous and exogenous covariates. Estimation of the posterior ordinate can be facilitated using the decomposition

$$\pi(\boldsymbol{\lambda}^*, \Omega^*, \boldsymbol{\delta}^*|\mathbf{y}) = \pi(\boldsymbol{\lambda}^*|\mathbf{y})\pi(\Omega^*|\mathbf{y}, \boldsymbol{\lambda}^*)\pi(\boldsymbol{\delta}^*|\mathbf{y}, \Omega^*, \boldsymbol{\lambda}^*).$$

Estimation of $\pi(\boldsymbol{\lambda}^*|\mathbf{y})$ is done by averaging the full conditional density with draws $\{\mathbf{y}^{*(g)}, \Omega^{(g)}\} \sim \pi(\mathbf{y}^*, \Omega|\mathbf{y})$ from the main MCMC run for $g = 1, \ldots, G$,

$$\pi(\boldsymbol{\lambda}^*|\mathbf{y}) \approx G^{-1} \sum_{g=1}^{G} \pi(\boldsymbol{\lambda}^*|\mathbf{y}, \mathbf{y}^{*(g)}, \Omega^{(g)}).$$

The next ordinate, $\pi(\Omega^*|\mathbf{y}, \boldsymbol{\lambda}^*)$, can be estimated using a reduced run to obtain

$$\pi(\Omega^*|\mathbf{y}, \boldsymbol{\lambda}^*) \approx G^{-1} \sum_{g=1}^{G} \pi(\Omega^*|\mathbf{y}, \boldsymbol{\lambda}^*, \mathbf{y}^{*(g)}).$$

The last ordinate, $\pi(\boldsymbol{\delta}^*|\mathbf{y}, \Omega^*, \boldsymbol{\lambda}^*)$, which is unique to the ordered probit setting, is estimated using the methods in Chib and Jeliazkov (2001).

The benefits of Bayesian model comparison go beyond dealing with issues of variable selection and model uncertainty. In this context, additional benefits include understanding the type of endogeneity and the dependence structure between a set of outcome variables. This allows a researcher to distinguish between several competing specifications and further investigate the relationships of interest. Emphasis is placed on Bayesian techniques because considering both specifications is a straightforward extension of the methodology. Gibbs sampling methods employed in discrete data models already generate the latent $\mathbf{y}^*$s in the data augmentation part of a sampler. Employing these draws as data involves minimal additional coding as discussed in the next section.

3. BANK CONTAGION

The first application addresses financial contagion in two ways. First, this application examines both latent and observed measures of a bank failure and how these determine nearby bank performance. The impact of a regional bank failure on bank health remains unclear in the existing literature. Calomiris and Mason (2003b) find that a nearby failure decreases the probability of survival for the remaining banks. However, there is also research on efficient bank runs (Freixas & Rochet, 2008), which finds positive effects stemming from nearby bank failures due to market competition. If an inefficient bank fails, its customers can go to the remaining banks in the market for deposits and lending, thereby benefiting the existing depository institutions. Second, the application evaluates different linkages for the spread of contagion. Channels for contagion have been found in both regional and correspondent[2] networks (Aharony & Swary, 1996; Kaufman, 1994; Richardson, 2007; Richardson & Troost, 2009). While both linkages may be present, financial regulators need to determine which channel is stronger in order to implement policies for restricting the spread of contagion and preventing bank runs. This paper addresses these issues and examines the consequences of bank failures during the 1930s by looking at town-wide failure rates and changes to correspondent networks.

The models for banks $i = 1,\ldots,n$, are those in Eqs. (1) and (2) where $\varepsilon_i \equiv (\varepsilon_{i1}, \varepsilon_{i2})' \sim N_2(0,\Omega)$ and $\Omega = \begin{pmatrix} 1 & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$. The models are characterized with two dependent variables in which $\mathbf{y}_i^* \equiv (y_{i1}^*, y_{i2}^*)'$ are the continuous latent data and $\mathbf{y}_i \equiv (y_{i1}, y_{i2})'$ are the corresponding discrete observed data. For the first outcome, the latent variables relate to the observed binary outcomes by $y_{i1} = 1\{y_{i1}^* > 0\}$ and

$$y_{i1} = \begin{cases} 0 & \text{No bank failure occurred nearby between 1929 and 1932} \\ 1 & \text{Bank failure occurred nearby between 1929 and 1932.} \end{cases}$$

The first outcome y_{i1} indicates whether or not a bank failure occurred between 1929 and 1932 in the town the subject bank does business. Ω incorporates the usual unit variance restriction in probit models, which is a normalization for identification. The second outcome y_{i2} measures a bank's performance in 1933, where there is point mass at 0 for banks that were suspended since 1932 and a continuous distribution with "loans and discounts" (hereafter referred to as LD) representing bank health. LD is chosen to measure a bank's performance following the literature on the credit crunch and its relation to economic activity (Bernanke, 1983; Calomiris & Mason, 2003a). The latent variables $\{y_{i2}^*\}$ relate to the observed censored outcomes by $y_{i2} = y_{i2}^* \cdot 1\{y_{i2}^* > 0\}$.

The endogenous covariate, y_{i1} in (1) or y_{i1}^* in (2), displays the impact of a nearby or regional bank failure on a bank's health and lending. Both the latent and observed specifications are easily motivated by hypotheses discussed in the existing literature. The latent counterpart of a regional bank failure can reveal unobserved factors that affect bank distress and profitability, such as corporate governance, risk behavior, and loss of confidence. Although an econometrician can observe whether a bank fails, additional information bankers have on the performance of their bank and business environment remains unobservable. In addition, existing evidence suggests that bank runs, which were the main mechanisms that caused bank failure (Bernanke, 1983), were often facilitated through "word-of-mouth" or information-based contagion (Park, 1991). For instance, if an individual's neighbor speculated about a pending bank failure, the individual is likely to withdraw deposits from his bank to avoid undue losses. Depositors lack financial information, resulting in withdrawal decisions based on the condition of the banking system as a whole (Park, 1991). Although researchers cannot measure the speculative nature of banking panics, the latent specification can act as a proxy for these factors. These arguments formally motivate the latent endogeneity model.

On the other hand, the literature notes that bank failures trigger panic (Chen, 1999). Despite speculation about bank health, depositors do not react until an indicator for failure is triggered. The literature also notes that publicizing the names of failed banks worsened remaining bank health. Additionally, the publication of the names of banks receiving financial assistance mitigated lender of last resort relief efforts (Butkiewicz, 1995). These observed outcomes, or triggers, support the observed endogeneity model. As mentioned previously, although a researcher may have an *a priori* expectation of the correct specification, it is hard to completely rule out the opposing approach. Therefore, a formal model comparison is necessary to better understand how a nearby failure affects bank performance and to ensure the employed specification accurately captures the interactions and decisions of banks during financial crises.

The data collected for this application are from the *Rand McNally Bankers' Directory*. This directory details balance sheets, correspondent relationships, and characteristics for all banks in a given state. Additional data are gathered from the 1930 U.S. census of agriculture, manufacturing and population, which describe the characteristics of the county and banks' business environment. The sample includes all banks operating in 1932 in Alabama, Arkansas, Michigan, Mississippi, and Tennessee for a total of 1,794 banks. These five states are considered because they provide variation across bank characteristics, size, Federal Reserve districts, and county characteristics. Table 1 presents descriptive statistics on the banks and average county characteristics for each state. For further information on the data set, see Vossmeyer (2014).

A key covariate of interest listed in Table 1 is Δ *Correspondents* – an indicator variable that takes the value 1 if a correspondent was removed from a bank's network and 0 otherwise. Recall that there are two linkages for contagion, regional which the nearby bank failure variable (y_{i1}) captures, and correspondent networks which this variable captures. Correspondent banks are usually designated in reserve cities of the Federal Reserve system and often provided smaller, local banks with liquidity (Richardson & Troost, 2009). Correspondent relationships between bigger and smaller banks built a structure for the Federal Reserve to influence nonmember institutions. However, this structure created pathways for contagion to spread. Therefore, controlling for it in the model is important and of interest to policy-makers in order to mitigate contagion through its many channels.

3.1. Estimation

The model is completed by specifying the prior distributions. For $\boldsymbol{\lambda} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$, $\pi(\boldsymbol{\lambda}) = \mathcal{N}(\boldsymbol{\lambda}|\mathbf{d}_0, \mathbf{D}_0)$, and $\pi(\Omega) \propto \mathcal{IW}(\nu_0, \mathbf{R}_0)1\{\omega_{11} = 1\}$, where the prior on Ω (inverse Wishart) is on the derived quantities that appear in Algorithm 1. The hyperparameters for the priors are selected using a training sample of 100 banks. A thorough sensitivity analysis is provided in Section 3.2. Algorithm 1 presents the Gibbs sampling and data augmentation methods to simulate the posterior distribution for the observed endogeneity specification.

Table 1. Financial Characteristics of the Banks in Each State in 1932 and County Characteristics.

Variable	Alabama	Arkansas	Michigan	Mississippi	Tennessee
Number of banks	250	278	638	235	393
Average age	24	22	30	25	25
Federal Reserve district	6	8	7, 9	6, 8	6, 8
Financial characteristics (avg. $1,000)					
Deposits	716	437	1750	528	765
Deposits/Liab. (*ratio*)	0.612	0.750	0.751	0.730	0.675
Loans & discounts (LD)	582	273	1203	356	649
Charters and memberships (counts)					
State bank	166	222	438	208	308
National bank	82	44	102	26	83
Amer Bk Ass'n (ABA)	160	188	370	171	185
Correspondents (averages)					
Total correspondents	2.6	2.4	2.8	2.9	2.4
Out of state corres.	1.5	1.4	1.5	2.5	1
Δ Correspondents	0.19	0.31	0.22	0.28	0.40
Market shares (averages)					
Liab./County Liab.	0.27	0.26	0.13	0.33	0.24
Liab./Town Liab.	0.68	0.75	0.17	0.74	0.69
Herfindahl index (HHI)	0.66	0.60	0.29	0.70	0.54
County characteristics (averages)					
No. of wholesale retailers	31.3	22.4	45.3	15.4	25.3
Cropland (×1,000 acres)	115.6	96.6	122.9	81.9	78.1
Town pop. (×1,000)	14.6	4.7	49.6	4.5	14.2

Algorithm 1. *MCMC Estimation Algorithm – Observed Specification.*

1. *Sample* $[\lambda|\mathbf{y}^*, \Omega] \sim N(\hat{\boldsymbol{d}}, \hat{\boldsymbol{D}})$, *where* $\hat{\boldsymbol{d}}$ *and* $\hat{\boldsymbol{D}}$ *are given by* $\hat{\boldsymbol{d}} = \hat{\boldsymbol{D}}\left(\boldsymbol{D}_0^{-1}\boldsymbol{d}_0 + \sum_{i=1}^n \boldsymbol{W}_i'\Omega^{-1}\mathbf{y}_i^*\right)$ *and* $\hat{\boldsymbol{D}} = \left(\boldsymbol{D}_0^{-1} + \sum_{i=1}^n \boldsymbol{W}_i'\Omega^{-1}\boldsymbol{W}_i\right)^{-1}$ *where*

$$\boldsymbol{W}_i = \begin{pmatrix} \mathbf{x}'_{i1} & 0 & 0 \\ 0 & \mathbf{x}'_{i2} & y_{i1} \end{pmatrix}.$$

2. *Sample* Ω *in a one-block, two-step procedure by drawing* $\omega_{22\cdot 1} \equiv \omega_{22} - \omega_{21}\omega_{11}^{-1}\omega_{12}$ *and* ω_{21}, *then reconstructing* Ω *from these quantities,*
   (a) $\omega_{22\cdot 1} \sim \mathcal{IW}(\nu_0 + n, Q_{22})$
   (b) $\omega_{21} \sim \mathcal{N}(Q_{11}^{-1}Q_{12}, \omega_{22\cdot 1}Q_{11}^{-1})$, *where*

$$\mathbf{Q} = \mathbf{R}_0 + \sum_{i=1}^{n} (\mathbf{y}_i^* - \mathbf{W}_i\boldsymbol{\lambda})(\mathbf{y}_i^* - \mathbf{W}_i\boldsymbol{\lambda})'.$$

and $\boldsymbol{Q}$ *is partitioned conformably with* Ω, *i.e.*,

$$\mathbf{Q} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}.$$

3. *For i=1,..., n, sample* $y_{i1}^* | y_{i2}^*, y_{i1}, \boldsymbol{\lambda}, \Omega \sim \mathcal{TN}_{\mathcal{A}_i}(\mu_{1|2}, V_{1|2})$, *where* $\mu_{1|2}$ *and* $V_{1|2}$ *are the usual conditional mean and conditional variance, respectively. If* $y_{i1} = 0$, $\mathcal{A}_i$ is $(-\infty, 0)$, *and if* $y_{i1} = 1$, $\mathcal{A}_i$ *is* $(0,\infty)$.
4. *For* $i{:}y_{i2} = 0$, *sample* $y_{i2}^* | y_{i1}^*, y_{i2}, \boldsymbol{\lambda}, \Omega \sim \mathcal{TN}_{\mathcal{A}_i}(\mu_{2|1}, V_{2|1})$, *where the region* $\mathcal{A}_i$ *is* $(-\infty, 0)$ *implied by the censoring of* y_{i2} in the truncated normal distribution.

This sampler can be easily adapted to handle the latent specification. The most convenient approach is to move to the reduced-form where the system in (2) can be rewritten as,

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ -\gamma_2 & 1 \end{pmatrix} \begin{pmatrix} y_{i1}^* \\ y_{i2}^* \end{pmatrix} &= \begin{pmatrix} \mathbf{x}'_{i1} & 0 \\ 0 & \mathbf{x}'_{i2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \\ \mathbf{A}\mathbf{y}_i^* &= \mathbf{X}_i\boldsymbol{\beta} + \varepsilon_i \\ \Leftrightarrow \mathbf{y}_i^* &= \mathbf{A}^{-1}X_i\boldsymbol{\beta} + \mathbf{A}_i^{-1}\varepsilon_i \\ \Leftrightarrow \mathbf{y}_i^* &= \mu_i + \nu_i \\ \nu_i &\sim \mathcal{N}(0, \mathbf{A}^{-1}\Omega\mathbf{A}^{-1'}). \end{aligned} \tag{6}$$

Simply change the data augmentation steps of the sampler (steps 3−4) to use the conditional mean and conditional variance from the reduced form, avoiding any additional computational burden brought on by considering both specifications. This is a straightforward adjustment because the $\mathbf{y}^*$s are already being generated. In the latent specification, they are generated and passed through the sampler as data, so $\boldsymbol{W}_i = \begin{pmatrix} \mathbf{x}'_{i1} & 0 & 0 \\ 0 & \mathbf{x}'_{i2} & y_{i1}^* \end{pmatrix}$ in steps 1 and 2 of Algorithm 1.

3. 2. Results

The results for the application are based on 11,000 MCMC draws with a burn-in of 1,000. The inefficiency factors for the parameters remain low with slightly higher values occurring for the parameters on the endogenous covariates and variables common to both equations. The point estimates for the exogenous covariates are similar across both specifications except in cases where the variable has a 95% credibility interval that includes 0. The following discussion covers the basic results for each equation, followed by the model comparison and sensitivity analysis.

Table 2 presents the posterior means and standard deviations for both specifications. The results indicate that a nearby bank failure (y_{i1}) positively affects lending for existing banks, which disagrees with some of the literature. However, unlike other studies, this paper is looking at a longer window of impact. Previous papers examine the immediate impact of a nearby failure, which is distress. Whereas, this paper finds that the long-run impact is positive, corroborating the research on efficient bank runs (Freixas & Rochet, 2008). This result supports the market competition hypothesis where failing banks leave additional depositors in the market as customers for the remaining banks. As a result, the failure of a nearby bank

Table 2. Banking Application – Posterior Means and Standard Deviations for the Bivariate System of Equations.

	Observed Endogeneity		Latent Endogeneity	
	Nearby Fail	Lending	Nearby Fail	Lending
Intercept	−5.023 (0.260)	−0.480 (0.137)	−6.514 (0.349)	−0.122 (0.138)
Financial characteristics				
Bank age		0.009 (0.002)		0.013 (0.002)
Lagged LD		0.134 (0.016)		0.107 (0.017)
National bank		0.367 (0.093)		0.488 (0.102)
Deposits/Liabilities		0.115 (0.094)		0.143 (0.111)
Δ Correspondents		−0.176 (0.081)		−0.138 (0.087)
Total correspondents	0.041 (0.029)	−0.014 (0.029)	−0.021 (0.039)	−0.032 (0.031)
County characteristics				
Town population	1.262 (1.008)	0.193 (0.327)	−0.658 (1.046)	−2.627 (0.540)
Wholesale retailers	0.001 (0.001)		0.001 (0.001)	
Acres of cropland	0.218 (0.481)		−0.025 (0.570)	
Liab./Town Liab.	2.273 (0.180)	−0.025 (0.119)	3.106 (0.238)	−0.152 (0.119)
Herfindahl index (HHI)	0.441 (0.111)		0.548 (0.131)	
No. of banks in town	1.000 (0.051)		1.426 (0.074)	
Fed. Dist 7	0.123 (0.110)	−0.619 (0.097)	0.181 (0.132)	−0.715 (0.097)
Fed. Dist 8	0.265 (0.101)	−0.027 (0.093)	0.280 (0.113)	−0.012 (0.092)
Fed. Dist 9	0.069 (0.207)	−0.239 (0.206)	0.021 (0.238)	−0.394 (0.213)
y_1		1.314 (0.090)		
y_1^*				0.122 (0.017)

strengthens the balance sheets for surviving depository institutions. The results also illustrate that the negative impacts of regional contagion diminish over time. Although the initial panic is not captured here, the contagion channel disappears and presents market benefits. The increased lending by the remaining banks in the town increases economic activity and restores confidence in the financial system.

The second channel for the spread of contagion is represented by the variable Δ *Correspondents*. The result for this covariate demonstrates that a reduction in a bank's correspondent network has a negative impact on bank lending. This result aligns with intuition because correspondent banks often provided short-term commercial paper to smaller banks and urged them to extend credit (Richardson & Troost, 2009). When a bank was removed or a failure occurred in these networks, it constrained the liquid assets available and, as a result, banks extended less credit. The marginal effect of a correspondent removal, averaged over both observations and MCMC draws, is approximately −0.037 and −0.051 for the latent and observed specifications, respectively. In other words, a reduction in a bank's correspondent network decreases lending by about \$370−\$510 for every \$10,000. This contagion channel presents moderate long-term negative effects. Policy-makers providing *ex post* liquidity following a downturn may want to focus relief efforts on banks experiencing failures across their correspondent networks as the negative effects may still be lingering.

Other results for the first equation display that the higher share of liability held at an individual bank (*Liab./Town Liab.*) and the higher the market concentration in a town (*HHI*), the more likely that town was to experience a bank failure between 1929−1932. In addition, relative to the sixth Federal Reserve district, the eighth district is more likely to experience bank failures, which accords well with Richardson and Troost's (2009) paper. The results for the second equation show that older banks and national banks have higher lending in 1933 relative to younger and nonmember institutions.

Model comparison results are presented in Table 3 and reveal that the data strongly support the latent endogeneity specification, where the full magnitude of y_{i1}^*, even values whose extent is driven by observed covariates and unobserved factors outside {0,1}, is relevant for bank lending. The marginal likelihood is nearly 50 points higher on the log scale, giving the observed specification a posterior model probability of approximately 0. The latent measure of a nearby bank failure captures unobserved aspects of bank profitability and the loss of confidence cultivating through these struggling local economies, which better explain bank lending. This result highlights an interesting point from Calomiris and Mason's (2003b) paper where they state, "Indicator variables are uninformative about the particular mechanism through which illiquidity and contagion produces a bank failure." Calomiris and Mason (2003b) further urge researchers to interpret indicators with caution when examining contagion as there may be evidence of missing fundamentals and loss of financial confidence. The issues these authors refer to can be mitigated by employing a latent measure for a nearby bank failure. The authors' analysis explains the strong support from the data for the latent specification and further corroborates the underlying hypothesis.

Without considering both modeling approaches, a researcher using observed endogeneity could misinterpret the relationship between regional failures and bank lending, and fundamentally misspecify financial panics. The nonlinear dichotomizing mechanism in which regional failures determine bank performance is inadequate. The latent variable approach encompasses the speculative nature of bank runs. One of the most documented and well understood features of banking crises is asymmetric information, which dates back to the original research by Bagehot (1873). Depositors, bankers, and central bankers lack credible information (Diamond & Dybvig, 1983; Gorton, 1985), resulting in speculative and fundamental bank runs. When an econometrician is examining the interactions involved in banking panics, asymmetric information should be apparent and the models being considered should reflect this feature.

Table 3 also displays the results for the variance–covariance matrix Ω. There is a negative correlation between the equations for a nearby bank failure and bank performance. After controlling for a number of balance sheet and county characteristics, the errors present a negative relationship, which implies there are harmful effects of a nearby bank failure not controlled for in the model. This result holds for both model specifications but there is a clear magnitude difference between the two. The observed

Table 3. Banking Application – Results for the Variance–Covariance Matrix and Marginal Likelihood Estimates.

	Observed Endogeneity	Latent Endogeneity
ω_{12}	−1.078 (0.053)	−0.305 (0.130)
ω_{22}	2.125 (0.087)	2.044 (0.088)
Ω_{corr}	$\begin{pmatrix} 1 & -0.740 \\ -0.740 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & -0.214 \\ -0.214 & 1 \end{pmatrix}$
Log-marginal lik. (numerical S.E.)	−1713.1 (0.423)	−1666.3 (0.245)

Table 4. Banking Application – Sensitivity Analysis for Different Training Sample Sizes.

Training Sample Size	Observed Log-Marg. Lik.	Latent Log-Marg. Lik.
No sample (priors centered at zero)	−1837.2	−1778.9
50	−1769.5	−1736.1
100	−1713.1	−1666.3
150	−1666.1	−1621.1
200	−1627.5	−1556.8

specification may overstate the relationship and the "true incidence of panic, since relevant fundamentals are likely omitted" from the model (Calomiris & Mason, 2003b).

The results discussed thus far employ a training sample prior of 100 banks. The sensitivity of the results to the hyperparameters is displayed in Table 4. The model rankings do not change across different training sample sizes. This indicates that the data speak loudly for the results and support the latent endogeneity model. Researchers interested in modeling and understanding decisions and relationships of banks in adverse macroeconomic conditions can employ latent variables to accommodate asymmetric information and to better capture the interactions between the outcomes of interest.

4. EDUCATION

The second application considers the impact of education on adult socioeconomic status. The return on schooling has been an ongoing area of research for empirical economists. Despite the great deal of attention focused on accommodating issues with survey responses, latent endogenous variables are lacking in this literature.

Education's impact on adult socioeconomic status is of particular interest because there are convincing arguments for both observed and latent endogeneity. Observed endogeneity is motivated by societal evaluations of education, which are generally looked at by crossing particular achievement thresholds or cut-points, e.g., high school degree and college degree. Without a college degree, individuals cannot apply for many jobs notwithstanding 15 years of schooling. Therefore, labor market outcomes are determined by these observed degree or threshold crossing indicators. Alternatively, a common issue in the analysis of returns to schooling is the inadequacy of measures for motivation and ability, which also explain socioeconomic status. Observed measures for ability have been proposed, such as standardized test scores. However, these data are often not available. Therefore, the latent representation of education can act as a proxy for unobservable characteristics linking education and socioeconomic outcomes. While the hypotheses for each approach are convincing, there have been no attempts to compare these models, which is necessary to ensure the most accurate specification is employed. Furthermore, knowing which approach is best supported by the data advances our understanding of how education determines socioeconomic status.

Studying education provides a unique opportunity for model comparison. Most education and socioeconomic outcomes are discretized by ordered categories. For instance, education categories can be defined by degree level and socioeconomic status can be categorized by income brackets. This ordinal data setting makes the observed specification in (1) invalid, and instead, the system in (3) can be employed. The specification for observed endogeneity for five education categories (defined shortly) and individuals $i=1,\ldots,n$ is:

$$
\begin{aligned}
y_{i1}^{*} &= \mathbf{x}'_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1} \\
y_{i2}^{*} &= \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + 1\{y_{i1}=2\}\gamma_{22} + 1\{y_{i1}=3\}\gamma_{23} + 1\{y_{i1}=4\}\gamma_{24} + 1\{y_{i1}=5\}\gamma_{25} + \varepsilon_{i2},
\end{aligned}
\tag{7}
$$

where $\varepsilon_i \equiv (\varepsilon_{i1}, \varepsilon_{i2})' \sim N_2(0,\Omega)$ and $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$. This specification differs from (1) because the endogenous covariate enters as a set of dummy variables for each category, whereas previously it entered as a single endogenous regressor.[3] The latent specification remains identical to the system in (2), which is now a more parsimonious model relative to the observed approach. It is important to note that the elements in Ω are left free. Location and scaling restrictions are accommodated by fixing two cutpoints. The different approaches for identification in multivariate ordered probit models are discussed in Jeliazkov et al. (2008).

The model is characterized by two dependent variables, where $\mathbf{y}_i^{*} \equiv (y_{i1}^{*}, y_{i2}^{*})'$ are the continuous latent data and $\mathbf{y}_i \equiv (y_{i1}, y_{i2})'$ are the corresponding discrete observed data. For equations $k=1,2$, the latent variables relate to the observed ordered outcomes by $y_{ik} = \sum_{j=1}^{J} 1\{y_{ik}^{*} > \alpha_{k,j-1}\}$ for J ordered alternatives where α_{kj} is a cut-point between the categories given by

$$y_{i1} = \begin{cases} 1 & \text{Less than high school} \\ 2 & \text{High school degree} \\ 3 & \text{Some college} \\ 4 & \text{College degree} \\ 5 & \text{Graduate education} \end{cases}$$

$$y_{i2} = \begin{cases} 1 & \text{Poverty line and below} \\ 2 & \text{Lower-middle class} \\ 3 & \text{Middle class and up.} \end{cases}$$

The outcome y_{i1} represents the amount of education an individual completes and y_{i2} represents an individual's socioeconomic status. The second outcome is measured by an income-to-needs ratio. Income is measured using the actual amount of total income, which is the sum of taxable income and transfer income. Needs is measured as the poverty threshold taken from the Census Bureau. These thresholds are based on family size and age of the household. An income-to-needs ratio below 1.3 indicates the poverty line and below, between 1.3 and 3 indicates lower-middle class, and above 3 represents the middle class and up.[4] The endogenous covariate y_{i1} represents the impact of education on adult socioeconomic status.

The data collected for this application are from the Panel Study of Income Dynamics (PSID). The sample includes 2,779 respondents from the 1999 survey. The data set contains information on childhood health, parental socioeconomic status, parental education, adult socioeconomic status, adult health, and educational attainment for individuals between the ages of 30 and 50. The year 1999 is selected because it features retrospective reports on childhood health. Table 5 offers descriptive statistics on the data and details the discretization for a number of variables.

4.1. Estimation

The model is completed by specifying the prior distributions,

$$\boldsymbol{\lambda} \sim \mathcal{N}(\boldsymbol{d}_0, \mathbf{D}_0)$$

Table 5. Descriptive Statistics for the Sample of 2,779 Respondents from the PSID.

Variable	Sample Proportion	Variable	Sample Proportion
Respondent education (Educ)		**Mother's education** (Meduc)	
<High school degree	0.17	<High school degree	0.33
High school degree	0.32	High school degree	0.47
Some college	0.25	Some college	0.09
College degree	0.17	College degree	0.08
Graduate school	0.09	Graduate school	0.03
Father's education (Feduc)		**Childhood health**	
<High school degree	0.39	Poor	0.16
High school degree	0.38	Average	0.38
Some college	0.07	Excellent	0.46
College degree	0.10	**Adult health**	
Graduate school	0.06	Poor	0.10
Adult socioeconomic status		Average	0.62
Low	0.13	Excellent	0.28
Medium	0.27	**Marital status**	
High	0.60	Single	0.16
Parental socioeconomic status (pSES)		Divorced	0.24
Low	0.25	Married	0.60
Medium	0.45	**Race**	
High	0.30	White/Asian	0.63
Debt		Non-white	0.37
Debt	0.55	**Employment**	
No debt	0.45	Employed	0.88
Sex		Unemployed	0.12
Male	0.76	**Age (average)**	40
Female	0.24		

$$\mathbf{\Omega} \sim \mathcal{IW}(\nu_0, \mathbf{R}_0).$$

The hyperparameters are selected using a training sample of 200 individuals. Algorithm 2 presents the sampling methods to simulate the posterior distribution for the observed endogeneity specification, which follow from Jeliazkov et al. (2008). Note, the cut-point parameters α_{kj} are transformed to ensure the ordering constraints, so

$$\delta_{kj} = \ln(\alpha_{kj} - \alpha_{k,j-1}),$$

where $2 \leq j \leq J - 1$ for equations $k = 1, 2$.

Algorithm 2 *MCMC Estimation Algorithm – Observed Specification*

1. *For each equation k, sample* $\boldsymbol{\delta}_k, \mathbf{y}_k^* | \mathbf{y}, \boldsymbol{\lambda}, \mathbf{y}_{\setminus k}^*$ *as follows*
   (a) *Sample* $\boldsymbol{\delta}_k | \mathbf{y}, \boldsymbol{\lambda}, \Omega, \mathbf{y}_{\setminus k}^*$ *using the Metropolis–Hastings algorithm*
   (b) *Sample* $y_{ik}^* | \mathbf{y}, \boldsymbol{\lambda}, \Omega, \mathbf{y}_{\setminus k}^* \sim \mathcal{TN}_{(\alpha_{k,j-1}, \alpha_{kj})}(\mu_{k|\setminus k}, V_{k|\setminus k})$ *for* $i = 1, \ldots, n$.
2. *Sample* $[\boldsymbol{\lambda} | \mathbf{y}^*, \Omega] \sim N(\hat{\boldsymbol{d}}, \hat{\boldsymbol{D}})$, *where* $\hat{\boldsymbol{d}}$ *and* $\hat{\boldsymbol{D}}$ *are given by* $\hat{\boldsymbol{d}} = \hat{\boldsymbol{D}}\left(\boldsymbol{D}_0^{-1}\boldsymbol{d}_0 + \sum_{i=1}^n \boldsymbol{W}_i'\Omega^{-1}\mathbf{y}_i^*\right)$ *and* $\hat{\boldsymbol{D}} = \left(\boldsymbol{D}_0^{-1} + \sum_{i=1}^n \boldsymbol{W}_i'\Omega^{-1}\boldsymbol{W}_i\right)^{-1}$.
3. *Sample* $\Omega \sim \mathcal{IW}(\nu_0 + n, \mathbf{R}_0 + (\mathbf{y}^* - \mathbf{W}\boldsymbol{\lambda})'(\mathbf{y}^* - \mathbf{W}\boldsymbol{\lambda}))$.

As a matter of notation, "$\setminus k$" is used to represent all elements in a set except the kth one. Estimation of the latent specification follows these steps closely, however, employs the reduced-form trick discussed in (6) for the data augmentation step in 1(b).

4.2. Results

The results are based on 11,000 MCMC draws with a burn-in of 1,000. Analysis of the sensitivity of the results to the training sample size is conducted as in the first application. The results again show no sensitivity to the training sample size and model rankings do not change for different hyperparameters. The inefficiency factors for the parameters remain low with the highest values (≈20) occurring for the parameters on the endogenous covariates in both specifications. The following discussion reviews the basic results for each equation, then the model comparison results.

Table 6 presents the posterior means and standard deviations for both specifications. The results from the first equation accord well with the existing literature on the determinants of educational attainment (Haveman & Wolfe, 1995). The results show that parental education and parental socioeconomic status play a positive role in educational attainment. Parents with more income are able to invest in their children's education with schooling supplies, tutors, and financial assistance in college. Furthermore, parents who themselves achieve a higher level of education often motivate their children to do the same. The results also indicate that whites complete more education relative to non-whites, and males complete more schooling relative to females.

The results of the second equation coincide well with intuition and what is often found in the literature on socioeconomic achievement. Parental socioeconomic status positively affects adult socioeconomic status and individuals who are married have a higher income-to-needs ratio relative to divorced individuals. An interesting result is the positive coefficient on the

Table 6. Education Application – Posterior Means and Standard Deviations for the Bivariate System of Equations.

	Observed Endogeneity		Latent Endogeneity	
	Education	SES	Education	SES
Intercept	−0.683 (0.593)	−1.706 (0.193)	−0.696 (0.597)	−1.353 (0.199)
Child health – Exc	−0.269 (0.586)		−0.265 (0.586)	
Child health – Avg	−0.429 (0.587)		−0.425 (0.586)	
Feduc – High school	0.188 (0.058)		0.189 (0.059)	
Feduc – Some college	0.616 (0.097)		0.613 (0.095)	
Feduc – College	0.840 (0.094)		0.835 (0.093)	
Feduc – Graduate	1.033 (0.114)		1.028 (0.115)	
Meduc – High school	0.197 (0.060)		0.204 (0.060)	
Meduc – Some college	0.499 (0.093)		0.498 (0.092)	
Meduc – College	0.534 (0.103)		0.531 (0.102)	
Meduc – Graduate	0.780 (0.147)		0.775 (0.147)	
pSES – High	0.339 (0.067)	0.073 (0.065)	0.351 (0.064)	0.035 (0.067)
pSES – Med	0.344 (0.053)	0.079 (0.058)	0.351 (0.060)	0.049 (0.059)
Single		−0.003 (0.064)		0.005 (0.064)
Married		0.154 (0.062)		0.168 (0.062)
Employed		0.634 (0.067)		0.637 (0.067)
Debt		0.202 (0.042)		0.196 (0.042)
Adult health – Exc		0.359 (0.080)		0.352 (0.080)
Adult health – Avg		0.244 (0.069)		0.245 (0.070)
Educ – High school		0.553 (0.086)		
Educ – Some college		0.813 (0.127)		
Educ – College		1.312 (0.172)		
Educ – Graduate		1.702 (0.241)		
Latent Educ (y_1^*)				0.424 (0.055)
White	0.344 (0.053)	0.416 (0.056)	0.347 (0.053)	0.389 (0.056)
Age	0.030 (0.004)	0.015 (0.004)	0.030 (0.004)	0.014 (0.004)
Male	0.101 (0.055)	0.359 (0.064)	0.101 (0.054)	0.358 (0.063)

debt variable. The debt variable is measured by summing all debt excluding debt from the purchase of a house. However, debt does not necessarily indicate financial distress, as wealthier individuals have more access to debt and may finance other purchases, including vehicles, businesses, and school loans. The results also indicate that health has a positive relationship with wealth. Healthy individuals are less likely to miss work due to illness or disability and are likely to be more productive.

The covariate of interest, education, positively impacts adult socioeconomic status in both specifications. Relative to no high school degree, the coefficient on each discrete category for additional schooling gets

Table 7. Education Application – Results for the Variance–Covariance Matrix and Marginal Likelihood Estimates.

	Observed Endogeneity	Latent Endogeneity
ω_{11}	1.258 (0.043)	1.263 (0.043)
ω_{12}	−0.168 (0.074)	−0.195 (0.072)
ω_{22}	0.739 (0.049)	0.744 (0.057)
Ω_{corr}	$\begin{pmatrix} 1 & -0.174 \\ -0.174 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & -0.200 \\ -0.200 & 1 \end{pmatrix}$
Log-marginal lik. (numerical S.E.)	−7864.0 (0.053)	−7862.9 (0.044)
$\Pr(\mathcal{M}_k \| y)$	0.333	0.667

incrementally larger in the observed specification. The results for Ω, presented in Table 7, demonstrate that after controlling for family background, health, and other demographics, there is a negative correlation between education and adult socioeconomic status. Although the direct effect of education on wealth is positive, there is a negative relationship between the errors. This result has been noted in the literature (Becker & Chiswich, 1966; Griliches, 1977), and a simple explanation for this is luck. The basic explanatory variables do not control for instances of good fortune.

Table 7 also presents the model comparison results. The marginal likelihoods and posterior model probabilities are very close for both specifications with a slight preference toward the latent endogeneity model. The latent specification is more parsimonious, with three less covariates than the observed endogeneity model. Further investigation into this relationship is necessary due to the lack of model preference. An approach for understanding these interactions is to separate the sample by age cohort. The intuition for this comes from recognizing that individuals' degree level may only be influential in obtaining their first few jobs. As the amount of time an individual is in the labor force increases, the individual accumulates work history, skills, and references, which mitigate the importance of degree level. Consider job postings for entry-level positions, the salary level is often listed as "competitive", whereas for more senior positions, "depending on experience" is a common listing. Following this intuition, both models are compared for two separate age cohorts. The results are displayed in Table 8.

The results align with the age group hypotheses. The data support the observed specification for the 30–35 age cohort. Alternatively, the data support the latent specification for the 40–50 age cohort. This is a major

Table 8. Education Application – Model Comparison and Marginal Likelihood Estimates for Different Age Cohorts.

	Ages 30–35		Ages 40–50	
	Observed	Latent	Observed	Latent
Observations	843	843	1489	1489
Log-marginal lik.	−2450.6	−2455.3	−4196.1	−4192.0
$\Pr(\mathcal{M}_k\|y)$	0.991	0.009	0.017	0.983

result because it displays how the dependence structure between educational attainment and adult socioeconomic status changes with age. For younger individuals, the primary way to signal intelligence and ability is through degree-level. As a result, society evaluates the amount of education completed by observed threshold crossing indicators, thus resulting in labor market outcomes for these individuals. On the other hand, for older individuals, the data support the underlying latent specification, capturing ability, work history, and other factors unobserved by the econometrician. The importance of education diminishes as other elements become more prominent and, eventually, better explain labor market outcomes and wealth. Additionally, the latent approach offers a unique opportunity for studies examining outcomes of older cohorts. Surveys often do not contain a comprehensive work history of the respondents, therefore, if these data are unavailable, the latent measure of education can offer some insight into these characteristics.

These results truly stress the importance of considering and comparing both specifications and contribute to the literature on returns to schooling by demonstrating the evolution in the dependence structure between education and socioeconomic status. Pathway models employed in labor economics attempt to capture life-cycle interactions and intergenerational transmissions of education, health, and wealth. The complexity of these models increases because these outcomes are dynamic and change over time. Future research can employ both observed and latent measures of these outcomes to better explain these relationships pertaining to age cohorts.

5. CONCLUDING REMARKS

This paper addresses an important but often overlooked issue, which is the proper specification of endogenous covariates. In multivariate discrete data models, endogeneity can be based on latent or observed data. Bayesian model comparison techniques can be employed to determine which approach is best supported by the data, thus increasing the understanding of the nature of endogeneity and the dependence structure between the relationships being modeled.

While most applied researchers have an *a priori* expectation of the correct model, it is extremely difficult to rule out hypotheses in support of the alternative approach, which is apparent from the two empirical applications considered in this paper. Therefore, model selection provides important insights that resolve competing hypotheses about the interactions of interest. The results from the first application show that the latent representation of a bank failure better explains regional financial contagion, relative to conditioning on its observed counterpart. The latent measure captures the speculative nature of banking panics and accommodates asymmetric information issues discussed in the literature, thus providing a more accurate model of banking crises. The results for the second application show that the observed representation of education better explains socioeconomic outcomes for younger cohorts. This dependence structure changes as individuals age and the latent measure of education becomes more meaningful. This paper employs a bivariate system of equations in the applications, however, these approaches are generalizable to a number of methodologies.

These results stress the importance of employing model selection techniques to distinguish between competing specifications. The issues discussed here are present in any multivariate discrete data setting and should be addressed in a number of applied literatures. Ignoring these techniques can cause a researcher to misinterpret the depicted relationships and misunderstand the nature of endogeneity.

NOTES

1. The cut-points α_{kj}, which are defined in the link function for ordinal data, are transformed such that $\delta_{kj} = \ln(\alpha_{kj} - \alpha_{k,j-1})$. A discussion of this transformation is offered in Section 4.1.

2. Correspondents were banks with ongoing relationships facilitated by deposits of funds (Richardson, 2007). These networks linked banks across the country and indicated the extent to which a bank was important within the national network of banking (Calomiris, Mason, Weidenmier, & Bobroff, 2013).

3. If y_{i1} entered directly, this would lead to a cardinal interpretation of the categories, which is incorrect.

4. The cut-point 1.3 is selected because it is the threshold for food stamps.

ACKNOWLEDGMENT

Special thanks to Ivan Jeliazkov for his invaluable guidance. I am also grateful to David Brownstone, Sean Dowsing, Ben Gillen, Dale Poirier, Arshad Rahman, Gary Richardson, Michael Sacks, and participants at the Advances in Econometrics Conference on Bayesian Model Comparison for their helpful comments.

REFERENCES

Aharony, J., & Swary, I. (1996). Additional evidence on the information-based contagion effects of bank failures. *Journal of Banking and Finance*, *20*, 57–69.

Bagehot, W. (1873). *Lombard street: A description of the money market*. London, NY: H.S. King.

Becker, G., & Chiswich, B. (1966). Education and the distribution of earnings. *American Economic Review*, *56*, 358–369.

Bernanke, B. S. (1983). Nonmonetary effects of the financial crisis in propagation of the great depression. *American Economic Review*, *73*, 257–276.

Butkiewicz, J. (1995). The impact of lender of last resort during the great depression: The case of the reconstruction finance corporation. *Explorations in Economic History*, *32*, 197–216.

Calomiris, C., & Mason, J. (2003a). Consequences of bank distress during the great depression. *American Economic Review*, *93*, 937–947.

Calomiris, C., & Mason, J. (2003b). Fundamentals, panics, and bank distress during the depression. *American Economic Review*, *93*, 1615–1645.

Calomiris, C. W., Mason, J. R., Weidenmier, M., & Bobroff, K. (2013). The effects of the reconstruction finance corporation assistance on Michigan's banks' survival in the 1930s. *Explorations in Economic History*, *50*, 525–547.

Chen, Y. (1999). The role of the first-come, first-served rule and information externalities. *Journal of Political Economy*, *107*, 946–968.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, *90*, 1313–1321.

Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, *96*, 270–281.

Diamond, D., & Dybvig, P. (1983). Bank runs, deposit insurance and liquidity. *Journal of Political Economy*, *91*, 401–419.

Freixas, X., & Rochet, J. C. (2008). *Microeconomics of banking*. Cambridge, MA: MIT Press.

Gorton, G. (1985). Bank suspension of convertibility. *Journal of Monetary Economics*, *15*, 177–193.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, *45*, 1–22.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, *33*, 1829–1878.

Jeliazkov, I., Graves, J., & Kutzbach, M. (2008). Fitting and comparison of models for multivariate ordinal outcomes. In *Bayesian econometrics* (Vol. 23, pp. 115–156). Advances in Econometrics. Bingley, UK: Emerald Group Publishing Limited.

Jeliazkov, I., & Lee, E. H. (2010). MCMC perspectives on simulated likelihood estimation. In *Maximum simulated likelihood* (Vol. 26, pp. 3–39). Advances in Econometrics. Bingley, UK: Emerald Group Publishing Limited.

Jeliazkov, I., & Rahman, M. A. (2012). Binary and ordinal data analysis in economics: Modeling and estimation. In X.-S. Yang (Ed.), *Mathematical modelling with multidisciplinary applications* (pp. 123–150). Hoboken, NJ: Wiley.

Kaufman, G. G. (1994). Banking contagion: A review of the theory and evidence. *Journal of Financial Services Research*, *8*, 123–150.

Li, M., & Tobias, J. (2014). Bayesian analysis of treatment effect models, In I. Jeliazkov & X.-S. Yang (Eds.), *Bayesian inference in the social sciences* (pp. 63–90). Hoboken, NJ: Wiley.

Maddala, G. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.

Mintz, O., Currim, I. S., & Jeliazkov, I. (2013). Information processing pattern and propensity to buy: An investigation of online point-of-purchase behavior. *Marketing Science*, *32*, 716–732.

Muthen, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117.

Park, S. (1991). Bank failure contagion in historical perspective. *Journal of Monetary Economics*, *28*, 271–286.

Richardson, G. (2007). Categories and causes of bank distress during the Great Depression, 1929–1933: The illiquidity versus insolvency debate revisited. *Explorations in Economic History*, *44*, 588–607.

Richardson, G., & Troost, W. (2009). Monetary intervention mitigated banking panics during the great depression: Quasi-experimental evidence from a federal reserve district boarder 1929–1933. *Journal of Political Economy*, *117*, 1031–1073.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*, 528–549.

van Hasselt, M. (2014). Bayesian analysis of sample selection models. In I. Jeliazkov & X.-S. Yang (Eds.), *Bayesian inference in the social sciences* (pp. 91–113). Hoboken, NJ: Wiley.

Vossmeyer, A. (2014). Treatment effects and informative missingness with an application to bank recapitalization programs. *American Economic Review*, *104*, 212–217.