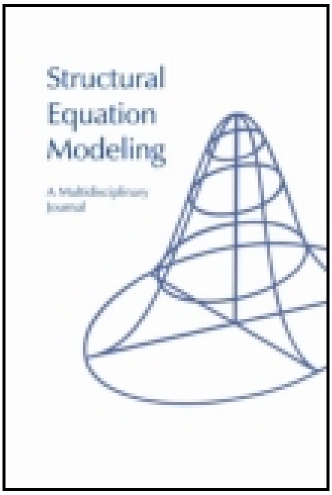

Structural Equation Modeling: A Multidisciplinary Journal

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hsem20

Estimating and Testing Mediation Effects with Censored Data

Lijuan Wang [a] & Zhiyong Zhang [a]

[a] University of Notre Dame

Published online: 07 Jan 2011.

Structural Equation Modeling, 18:18–34, 2011

ISSN: 1070-5511 print/1532-8007 online
DOI: 10.1080/10705511.2011.534324

Psychology Press
Taylor & Francis Group

Estimating and Testing Mediation Effects with Censored Data

Lijuan Wang and Zhiyong Zhang
University of Notre Dame

This study investigated influences of censored data on mediation analysis. Mediation effect estimates can be biased and inefficient with censoring on any one of the input, mediation, and output variables. A Bayesian Tobit approach was introduced to estimate and test mediation effects with censored data. Simulation results showed that the Bayesian Tobit approach can be used to deal with censoring in estimating and testing mediation effects under certain conditions. The proposed method was illustrated by analyzing an empirical data set with more than 50% of censored data on the output variable. Some guidelines for applying the proposed method were discussed.

In health and social science research, researchers sometimes observe censored data instead of complete data. In health research, Type I censoring (Klein & Moeschberger, 2003) occurs when participants in a clinical trial or a longitudinal study have periodic follow-up and patients' event time fall in an interval (L_i, R_i), where L_i is the left endpoint and R_i is the right endpoint of the censoring interval). When L_i equals $-\infty$ and R_i equals $+\infty$, the data are not censored. When L_i equals $-\infty$ and R_i is a finite number, the data are right censored. When R_i equals $+\infty$ and L_i is a finite number, the data are left censored. Otherwise, the data are interval censored. For example, in a longitudinal life span study investigating Alzheimer's disease (AD), researchers are interested in the age of initial occurrence of AD. For the participants who have their initial occurrence of AD during the study, researchers are able to observe the values of the response variable. However, for the participants who have had AD and do not know their ages of initial occurrence before entering the study, the observed values of the response variable are recorded as the ages when they enter the study. Furthermore, for those who do not have AD before or during the study, their observed values are recorded as the ages at the end of the study. Therefore, the age variable can be correctly observed only for a part of the participants. For the other participants, we have limited information on the response variable.

In educational and psychological measurement settings, we can also observe censored data. Right censored data or left censored data are also called ceiling data or floor data in testing settings. For example, when the test is relatively too easy or too difficult, the true abilities or traits of some individuals might not be measured accurately if the participants answer all items

Correspondence should be addressed to Lijuan Wang, 111 Haggar Hall, Department of Psychology, University of Notre Dame, South Bend, IN 46556, USA. E-mail: lwang4@nd.edu

correctly or incorrectly due to a low ceiling or a high floor threshold. In this case, the true abilities of those individuals who obtain maximum or minimum scores cannot be determined. Uttl (2005) provided a comprehensive discussion of severe ceiling effects in widely used memory tests, such as the verbal paired associates and word list tests from the Wechsler Memory Scales (Wechsler, 1945, 1987, 1997), the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964), and the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987). Among the adverse effects of low ceilings mentioned by Uttl were underestimated means and standard deviations and attenuated reliability and validity. It was also found that parameter estimates from regular regression analysis or structural equation models are biased and inefficient with censored data (e.g., Brown, 1992; Muthén, 1989, 1990; Tobin, 1958).

Different approaches have been proposed and applied to deal with censored data. For example, Muthén (1989) developed and applied the Tobit approach to analyze censored data for a factor analysis model in two steps. The Tobit correlation matrix was estimated in the first step, and the factor loadings and variances were estimated based on the Tobit correlation matrix in the second step. Brown (1992) compared different approaches to estimate a path model with latent variables for censored data via a simulation study and found that the procedure in Muthén (1989) provided relatively better results. Wang, Zhang, McArdle, and Salthouse (2008) introduced and applied the Tobit growth curve model to analyze longitudinal ceiling data and found the results were more accurate and precise than the regular growth curve model for longitudinal ceiling data.

In this study, we investigate potential influences of censored data on mediation analysis that is widely used in social and behavioral sciences. We introduce a Tobit mediation model to deal with censored data in mediation analysis. In the following, we first give a brief review of mediation analysis and hypothesis testings of mediation effects. Then we present the Tobit mediation model and Bayesian estimation method for mediation analysis with censored data. After that, the influences of censoring on estimating and testing mediation effects are investigated via simulations. The accuracy, precision, and power of the Bayesian Tobit mediation approach on estimating and testing mediation effects are also investigated via simulations. Finally, the application of the Bayesian Tobit mediation model is demonstrated using a real cognitive aging research example.

MEDIATION ANALYSIS AND HYPOTHESIS TESTING

Mediation analysis has been widely used in psychological and behavioral research to develop theories of whether there exists a third variable M that accounts for the relationship between an input variable X and an output variable Y (Baron & Kenny, 1986; Judd & Kenny, 1981a, 1981b; MacKinnon, Fairchild, & Fritz, 2007; Shrout & Bolger, 2002). The most widely applied mediation model perhaps is the observed three-variable model or the observed single-mediator model displayed in Figure 1. In Figure 1, Y, X, and M represent the dependent or output variable, the independent or input variable, and the mediation variable, respectively. e_M and e_Y are residuals with variances σ^2_{eM} and σ^2_{eY}. The mediation models can be expressed by two regression equations:

$$\begin{cases} Y = i_1 + c'X + bM + e_Y \\ M = i_2 + aX + e_M \end{cases}, \tag{1}$$

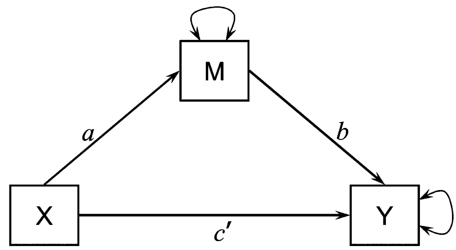

FIGURE 1 A path diagram for a simple mediation model.

where a represents the relationship between M and X, b represents the relationship between M and Y after controlling the effect of X, and c' represents the relationship between X and Y after controlling the effect of M. c' is also called the direct effect of X on Y and ab is called the indirect effect of X on Y through M. The implied model without the mediation effect is $Y = i_0 + cX + e_{Y0}$.

When mediation effects have occurred, the indirect effect, ab, or the difference in the direct effects, $c-c'$, should be significantly different from 0 (e.g., Baron & Kenny, 1986; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Shrout & Bolger, 2002; Sobel, 1982). If c' is not significantly different from 0, we call the effect a full mediation effect. Otherwise, we call it a partial mediation effect. In this study, we focus on the product ab for the test of mediation effect ($H_0 : ab = 0$) because it is the most widely used one.

For statistically testing the mediation effect, two major approaches have been used. The first approach, which can be viewed as a "single sample" method, is based on large-sample normal approximations. MacKinnon et al. (2002) gave a thorough review of 14 different single sample tests and compared their performance via a simulation study. Among the tests on the product of coefficients, the simulation results showed that the MacKinnon et al.'s $z' = ab/\sqrt{a^2\sigma_b^2 + b^2\sigma_a^2}$ test based on the empirical distribution of ab/σ_{ab} has the greatest power when both a and b are nonzero and has the most accurate Type I error rates when both a and b are zero. Based on the simulation results, the widely used Sobel (1982) first-order test $z = ab/\sqrt{a^2\sigma_b^2 + b^2\sigma_a^2}$ has relatively low power rates and Type I error rates. The second approach is based on the bootstrap resampling procedure (e.g., Bollen & Stine, 1990; Efron, 1979, 1987; Preacher & Hayes, 2004). The resampling methods do not require the large sample size assumption and could be more accurate and more powerful than the single sample method under certain conditions such as studies with small size, skewed outcome problems, or both (e.g., MacKinnon et al., 2007; MacKinnon, Lockwood, & Williams, 2004; Shrout & Bolger, 2002; Zhang & Wang, 2008).

THE TOBIT MEDIATION MODEL AND PARAMETER ESTIMATION

Suppose all three variables in the simple mediation model are possible to be interval censored with intervals ranging from L_X to R_X, L_M to R_M, and L_Y to R_Y for variables X, M, and Y, respectively. Let x^*, m^*, and y^* be the latent true scores and x, m, and y be the observed scores.

Then, we have,

$$x = \begin{cases} L_x & \text{if } x^* \leq L_x \\ x^* & \text{if } L_x < x^* < R_x \\ R_x & \text{if } x^* \geq R_x \end{cases},$$

$$m = \begin{cases} L_m & \text{if } m^* \leq L_m \\ m^* & \text{if } L_m < m^* < R_m \\ R_m & \text{if } m^* \geq R_m \end{cases},$$

and

$$y = \begin{cases} L_y & \text{if } y^* \leq L_y \\ y^* & \text{if } L_y < y^* < R_y \\ R_y & \text{if } y^* \geq R_y \end{cases}$$

For any of the three variables, when $L = -\infty$ and $R = +\infty$, the variable is not censored and thus we can observe its true scores. When $L = -\infty$ and $R \neq +\infty$, the variable is right censored and some true scores cannot be observed but are recorded as R. When $L \neq -\infty$ and $R = +\infty$, the variable is left censored and some scores are recorded as L.

Let $\mu = (\mu_X, \mu_M, \mu_Y)$ and

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XM} & \sigma_{XY} \\ \sigma_{MX} & \sigma_M^2 & \sigma_{MY} \\ \sigma_{YX} & \sigma_{YM} & \sigma_Y^2 \end{pmatrix}$$

denote the population mean vector and covariance matrix of the three variables. Let $\bar{Z} = (\bar{X}, \bar{M}, \bar{Y})'$ and

$$S = \begin{pmatrix} s_X^2 & s_{XM} & s_{XY} \\ s_{MX} & s_M^2 & s_{MY} \\ s_{YX} & s_{YM} & s_Y^2 \end{pmatrix}$$

denote the observed mean vector and the observed covariance matrix. Because of censoring, we can only observe the true scores when they are within the censoring intervals. Thus, the observed mean vector ($\bar{Z}$) and covariance matrix (S) are different from the population mean vector (μ) and covariance matrix (Σ) not only because of random sampling errors but also because of censoring errors. Therefore, the regular mediation analysis can lead to biased estimates ($\hat{a} = s_{XM}/s_X^2$ and $\hat{b} = (s_{MY}s_X^2 - s_{XM}s_{XY})/(s_X^2 s_M^2 - s_{XM}^2)$) and incorrect test results of mediation effects.

To deal with censoring problems in mediation analysis, we propose to use the Tobit mediation model. Instead of modeling the observed variables X, Y, and M directly, we model the latent variables X^*, Y^*, and M^* using the following model

$$\begin{cases} Y^* = i_1 + c'X^* + bM^* + e_{Y*} \\ M^* = i_2 + aX^* + e_{M*} \end{cases}, \tag{2}$$

where the scores for X^*, M^*, and Y^* are x^*, m^*, and y^*. Because x^*, m^*, and y^* are not available if they are censored, the estimation methods for the regular mediation model are not applicable. However, the maximum likelihood (MLE) Tobit and Bayesian Tobit approaches can be applied to estimate the model. In this study, we focus on the Bayesian approach because it has advantages over the MLE method in estimating and testing the Tobit mediation model.

The MLE Tobit approach is a multiple-stage method. In the first step, the means and variances of each latent variable (X^*, Y^*, and M^*) and the covariances of three pairs of latent variables ($X^*, Y^*; X^*, M^*; M^*, Y^*$) are estimated. The model parameters utilizing the estimated mean vector and covariance matrix are then estimated in the second step. In the third step, with estimates and standard errors for both a and b, we can apply the single sample method such as z' and z or the bootstrap resampling methods to conduct statistical inference on the mediation effect ab. However, the Bayesian Tobit approach can deal with censored data, estimating model parameters a and b, and estimating the mediation effect ab in a single step. In addition, using the Bayesian approach, the empirical distribution of ab can be obtained directly. Overall, the Bayesian method can deal with censored data and test the mediation effect based on the product parameter that might have a nonnormal distribution simultaneously and seamlessly.

To apply the Bayesian approach, we first need to specify the prior distributions for the model parameters. In this study, we use the normal priors for the regression coefficients and the inverse gamma priors for the variance parameters (see Appendix A). Based on these prior specifications, the conditional posterior distributions are obtained and provided in Appendix A. Using the conditional posterior distributions, the Gibbs sampling procedure (Chib, 1992) can be applied to generate a Markov chain for each parameter and the mediation effect in the Tobit mediation model.

Let $x_o, m_{o,} y_o$ denote the observed uncensored data and x_c, m_c, y_c denote the observed censored data. Furthermore, let x_c^*, m_c^*, y_c^* denote the true scores that are not available in the observed data set but can be sampled from the conditional posterior distributions of latent variables X^*, M^*, Y^*. Therefore, estimates of x^*, m^*, y^* can be obtained by augmentating the observed uncensored data $(x_o, m_{o,} y_o)$ and the underlying true scores (x_c^*, m_c^*, y_c^*) for the observed censored data. The Bayesian estimation for the Tobit mediation model can be implemented with the following steps:

1. Obtain the ordinary least squares estimates for the parameters including $u_X, \sigma_X^2, i_1, c', b, i_2, a, \sigma_{eY}^2$, and σ_{eM}^2 from the observed data $(x_o, m_{o,} y_o)$. These estimates are then used as the initial values for the Gibbs sampling algorithm. Provide initial values to x_c^*, m_c^*, y_c^* so that we have initial values for x^*, m^*, y^*.
2. Sample the true scores for the observed censored data.
   a. Sample $x_c^* | \mu_X, \sigma_X^2, a, b, c', i_1, i_2, \sigma_{eM}^2, \sigma_{eY}^2, y^*, m^*$. Generate $x_{c,i}^*$ from the truncated normal distribution $TNX_{(R_X, +\infty)}$ if the observed datum is censored and $x_{c,i}$ is equal to R_X or $TNX_{(-\infty, L_X)}$ if the observed datum is censored and $x_{c,i}$ is equal to L_X.

b. Sample $m_c^*|a, b, c', i_1, i_2, \sigma_{eM}^2, \sigma_{eY}^2, y^*, x^*$. Generate $m_{c,i}^*$ from the truncated normal distribution $TNM_{(R_M,+\infty)}$ if the observed datum is censored and $m_{c,i}$ is equal to R_M or $TNM_{(-\infty,L_M)}$ if the observed datum is censored and $m_{c,i}$ is equal to L_M.
c. Sample $y_c^*|i_1, c', b, \sigma_{eY}^2, x^*, m^*$. Generate $y_{c,i}^*$ from the truncated normal distribution $TNY_{(R_Y,+\infty)}$ if the observed datum is censored and $y_{c,i}$ is equal to R_Y or $TNY_{(-\infty,L_Y)}$ if the observed datum is censored and $y_{c,i}$ is equal to L_Y.

3. Sample the model parameters:
   a. Sample μ_X from the conditional posterior distribution $\mu_X|\sigma_X^2, x^*$.
   b. Sample σ_X^2 from the conditional posterior distribution $\sigma_X^2|\mu_X, x^*$.
   c. Sample a from the conditional posterior distribution $a|x^*, m^*, i_2, \sigma_{eM}^2$.
   d. Sample b from the conditional posterior distribution $b|x^*, m^*, y^*, i_1, c', \sigma_{eY}^2$.
   e. Sample c' from the conditional posterior distribution $c'|x^*, m^*, y^*, i_1, b, \sigma_{eY}^2$.
   f. Sample i_1 from the conditional posterior distribution $i_1|x^*, m^*, y^*, c', b, \sigma_{eY}^2$.
   g. Sample i_2 from the conditional posterior distribution $i_2|x^*, m^*, a, \sigma_{eM}^2$.
   h. Sample σ_{eM}^2 from the conditional posterior distribution $\sigma_{eM}^2|x^*, m^*, a, i_2$.
   i. Sample σ_{eY}^2 from the conditional posterior distribution $\sigma_{eY}^2|x^*, m^*, y^*, c', b, i_1$.
   j. Calculate the mediation effect ab.
4. Repeat Steps 2 and 3 until convergence and desired precision of the parameter estimates are met.

Step 2 is implemented for data imputation. Step 3 is included for sampling the unknown parameters from the conditional posterior distributions with augmented data. When there are no censored data, Step 2 can be skipped. In this case, we call the analysis the Bayesian regular mediation analysis. After a sufficient number of iterations, the burn-in period, the resulting samples from the after burn-in iterations can be viewed from the joint distributions of the model parameters (Geman & Geman, 1984). To check the convergence of the Gibbs sampling, we use both the graphical method by checking the overall pattern of the trace plot of Markov chains and the Geweke statistic (Geweke, 1992). If the trace plot appears stationary and the Geweke statistic is between −1.96 and 1.96, convergence is achieved.

With the converged samples from the conditional posterior distributions, we can directly obtain the empirical distribution of the product ab and then conduct statistical inference for the mediation effect based on the empirical distribution. For example, the point estimate can be obtained using the mean or the median and the standard error can be obtained using the standard deviation of the converged samples. The 95% confidence interval can be obtained based on the empirical 2.5% and 97.5% percentiles. Because the distribution of the product parameter might not be symmetric, the confidence interval constructed in this way might not be symmetric.

INFLUENCES OF CENSORING ON MEDIATION ANALYSIS

Intuitively, censoring might influence mediation analysis. In this section, we investigate possible influences of censoring on mediation analysis more concretely via a simulation study. In the simulation, the noncensored raw data were first generated with a median effect size ($a = 0.39$, $b = 0.39$) and a sample size of 200. Then the censored data were generated under different situations. Specifically, two types of censoring (interval censoring and left censoring) and three conditions of censored variables (the input variable is censored, the mediation variable is censored, and the outcome variable is censored) were considered. Furthermore, four different proportions of censored data (0%, 25%, 50%, and 75%) were also considered in generating the censored data. The number of replications is 1,000.

The simulated data were analyzed by the described Bayesian procedures in the preceding section without Step 2 (Bayesian regular mediation analysis) by assuming the censored data were the true data to investigate the influences of censoring on estimating and testing mediation effects. For the prior distributions of the model parameters, uninformative priors were used. For example, for all the regression coefficients, a normal distribution prior with mean 0 and variance $1.0E+6$ was used. For the variance or residual variance parameters, an inverse gamma distribution with a shape parameter of 0.001 and a scale parameter of 0.001 was used. A total of 10,000 burn-in iterations and 20,000 after burn-in iterations were used to generate the Markov chains for the model parameters. The Gibbs sampling is implemented in WinBUGS and the WinBUGS codes are contained in Appendix B. The simulations were conducted by using the SAS Macros developed by Zhang, McArdle, Wang, and Hamagami (2008).

By checking the Geweke statistic of each replication across different situations, we found that at least 99% of the 1,000 replications obtained convergent results. Results from those convergent replications are summarized in Table 1. In Table 1, we present five statistics on the mediation effect ab. First, point estimates of mediation effects were calculated by averaging the point estimates from different replications. Second, a measure of the bias in the mediation effect estimate, the relative bias of estimate ($RBE = \frac{\sum \hat{a}\hat{b}/R - .39\times.39}{.39\times.39}$), was calculated from dividing

TABLE 1
Influences of Censored Data on Estimating and Testing Mediation Effects ($n = 200$)

	0%	*M,I*	*M,L*	*X,I*	*X,L*	*Y,I*	*Y,L*
Proportion: 25%							
Estimate	0.152	0.140	0.134	0.188	0.182	0.114	0.114
RBE	0.2%	−8.2%	−11.6%	23.8%	19.8%	−24.8%	−24.9%
SE1	0.041	0.038	0.038	0.050	0.049	0.031	0.031
SE2	0.040	0.036	0.036	0.050	0.048	0.029	0.030
CP	0.954	0.945	0.928	0.905	0.923	0.780	0.767
Proportion: 50%							
Estimate	0.152	0.124	0.107	0.256	0.224	0.076	0.076
RBE	0.2%	−18.6%	−29.4%	68.3%	47.4%	−50.3%	−50.0%
SE1	0.041	0.037	0.035	0.072	0.066	0.021	0.022
SE2	0.040	0.034	0.033	0.070	0.062	0.020	0.022
CP	0.954	0.894	0.744	0.673	0.835	0.105	0.174
Proportion: 75%							
Estimate	0.152	0.109	0.070	0.449	0.300	0.038	0.038
RBE	0.2%	−28.7%	−54.2%	195.0%	97.6%	−74.9%	−74.9%
SE1	0.041	0.034	0.028	0.128	0.099	0.011	0.013
SE2	0.040	0.031	0.027	0.126	0.098	0.010	0.014
CP	0.954	0.757	0.289	0.223	0.676	0.000	0.000

Note. M = censored on the mediator variable; X = censored on the input variable; Y = censored on the output variable; I = interval censored; L = left censored; CP = coverage probability; *RBE* = relative bias of estimate; *SE1* = mean of the standard error estimates of all the replications; *SE2* = standard deviation of the parameter estimates of all the replications. The 0% column contains the results from the data with 0% censoring.

the difference between the mean of the parameter estimates from different replications and the true value by the true value. Third, precision of mediation effect estimates was evaluated by both the mean of the standard error estimates from different replications ($SE1$) and the standard deviation of the parameter estimates ($SE2$). The standard deviation of the parameter estimates ($SE2$) from the condition without any censored data can be viewed as the "true" standard error. Finally, the coverage probabilities based on 95% confidence levels were also obtained.

From Table 1, when there were no censored data (0% censoring), the Bayesian Tobit mediation approach can accurately estimate the mediation effect. The RBE value is 0.2% and the coverage probability is 95.4%. By comparing the average standard error estimate ($SE1 = 0.041$) and the standard deviation of parameter estimates ($SE2 = 0.040$), one can see that the standard errors were also accurately estimated. However, when there are censored data on any one of the three variables, the mediation effects were either underestimated or overestimated, with the RBE values ranging from −74.9% to +195% under different situations. For example, when the output variable was left censored with 50% censored data, the mediation effect was underestimated and the RBE value was −50%. When the input variable was left censored with 50% censored data, the mediation effect was overestimated and the RBE value was 47.4%. In terms of the coverage probability, the coverage probabilities based on the 95% confidence intervals were all smaller than 95% and ranged from 0% to 94.5%. By comparing the coverage probabilities across columns, we can find that censoring on the output variable has more influences on recovering the true values than censoring on the input or mediator variables. In addition, with larger proportions of censored data, the biases (RBE) were larger and the coverage probabilities were lower. For example, when the output variable was censored with 50% censored data, the coverage probability was lower than 20%. When the censoring proportion increased to 75%, the coverage probability was 0%. For the standard error estimates, the estimates were underestimated or overestimated with censored data and the relative biases ranged from −72.5% to 220%.

PERFORMANCE OF THE BAYESIAN TOBIT MEDIATION APPROACH

To examine the performance of the Bayesian Tobit mediation approach on testing and estimating mediation effects with censored data, the simulated censored data were analyzed by the Bayesian procedure with all the steps from 1 to 4 described earlier. The results from the simulation with a sample size of 200 are provided in Table 2. From Table 2, one can see that the proposed method can estimate the true mediation effects accurately even when the proportion of censored data was as large as 75%. For the situations with censoring on either the mediator variable or the input variable, the maximum RBE value was less than 1.5%. When the output variable was censored, the maximum RBE value was less than 8.6% with 75% of censored data. Furthermore, the coverage probabilities were all close to 95% and ranged from 94.9% to 96.4% regardless of different censoring types and different censoring proportions. However, with a higher proportion of censored data, the lengths of the confidence intervals were greater than the true length (0.158). For instance, the lengths were 0.175 or 0.212 when the output variable was left censored with 50% or 75% of censored data. For the standard errors, with 25% or 50% of censored data, the standard error estimates were slightly larger

TABLE 2
Results from the Bayesian Tobit Approach on Estimating and Testing Mediation Effects with Censored Data ($n = 200$)

	0%	*M,I*	*M,L*	*X,I*	*X,L*	*Y,I*	*Y,L*
Proportion: 25%							
Estimate	0.152	0.153	0.153	0.153	0.153	0.154	0.154
RBE	0.2%	0.5%	0.4%	0.3%	0.3%	1.4%	0.9%
SE1	0.041	0.041	0.042	0.041	0.041	0.042	0.042
SE2	0.040	0.041	0.040	0.041	0.040	0.041	0.041
Lower percentile	0.081	0.079	0.078	0.080	0.079	0.080	0.080
Upper percentile	0.239	0.241	0.242	0.240	0.241	0.243	0.243
CP	0.954	0.958	0.959	0.952	0.954	0.950	0.952
Proportion: 50%							
Estimate	0.152	0.153	0.153	0.152	0.152	0.157	0.155
RBE	0.2%	0.6%	0.4%	0.1%	0.2%	3.1%	2.1%
SE1	0.041	0.044	0.045	0.043	0.044	0.045	0.045
SE2	0.040	0.042	0.043	0.042	0.042	0.044	0.044
Lower percentile	0.081	0.075	0.072	0.077	0.075	0.079	0.077
Upper percentile	0.239	0.246	0.250	0.244	0.246	0.253	0.252
CP	0.954	0.957	0.962	0.953	0.950	0.952	0.956
Proportion: 75%							
Estimate	0.152	0.154	0.154	0.154	0.154	0.165	0.161
RBE	0.2%	1.4%	1.0%	1.1%	0.9%	8.6%	6.1%
SE1	0.041	0.047	0.054	0.047	0.050	0.052	0.054
SE2	0.040	0.045	0.053	0.046	0.049	0.052	0.053
Lower percentile	0.081	0.071	0.058	0.073	0.066	0.077	0.069
Upper percentile	0.239	0.255	0.270	0.257	0.263	0.282	0.281
CP	0.954	0.964	0.949	0.955	0.949	0.953	0.954

Note. M = censored on the mediator variable; X = censored on the input variable; Y = censored on the output variable; I = interval censored; L = left censored; CP = coverage probability; *RBE* = relative bias of estimate; *SE1* = mean of the standard error estimates of all the replications; *SE2* = standard deviation of the parameter estimates of all the replications. The 0% column contains the results from the data with 0% censoring.

than the true values. With 75% of censored data, the standard error estimates were about 18% to 35% larger than the true values. The estimated power was all 1 when the sample size was 200. From the preceding results, we can find that (a) the Tobit mediation analysis can provide more accurate and efficient estimates than the regular mediation analysis for censored data, and (b) compared to the complete data situation (0% of censored data), the mediation effect estimates are less efficient with censored data.

To further examine when the proposed approach might fail to recover true mediation effects, simulations were conducted with smaller sample sizes of 50 and 100 and the results are summarized in Tables 3 and 4. As one can see in Table 3, when the sample size was 100 and the proportion of censoring was 25% or 50%, the true mediation effects can be estimated well with the maximum RBE value of 6.8% from the situation of 50% interval censoring on the output variable. However, with 75% of censored data on the output variable, the RBE values became 19.3% and 14.4% for the situations of interval censoring and left censoring on the output variable. Furthermore, with 50% or 75% of censored data, the estimated power was smaller than the power without censoring and the estimated standard errors were larger than the true standard errors.

TABLE 3
Results from the Bayesian Tobit Approach on Estimating and Testing Mediation Effects with Censored Data ($n = 100$)

	0%	*M,I*	*M,L*	*X,I*	*X,L*	*Y,I*	*Y,L*
Proportion: 25%							
Estimate	0.152	0.154	0.153	0.153	0.153	0.156	0.155
RBE	0.0%	1.0%	0.9%	0.5%	0.3%	2.7%	1.9%
SE1	0.059	0.061	0.061	0.060	0.060	0.062	0.062
SE2	0.056	0.059	0.059	0.058	0.058	0.059	0.059
Lower percentile	0.050	0.049	0.048	0.050	0.049	0.051	0.050
Upper percentile	0.280	0.286	0.287	0.284	0.285	0.292	0.290
CP	0.957	0.955	0.952	0.952	0.949	0.954	0.961
Power	0.935	0.923	0.911	0.926	0.922	0.927	0.922
Proportion: 50%							
Estimate	0.152	0.155	0.153	0.152	0.153	0.162	0.159
RBE	0.0%	1.8%	0.8%	0.2%	0.4%	6.8%	4.7%
SE1	0.059	0.064	0.067	0.063	0.064	0.068	0.068
SE2	0.056	0.063	0.063	0.060	0.061	0.066	0.065
Lower percentile	0.050	0.044	0.039	0.046	0.043	0.049	0.046
Upper percentile	0.280	0.295	0.299	0.291	0.294	0.314	0.310
CP	0.957	0.949	0.959	0.955	0.958	0.955	0.955
Power	0.935	0.896	0.848	0.905	0.885	0.913	0.890
Proportion: 75%							
Estimate	0.152	0.157	0.155	0.155	0.155	0.181	0.174
RBE	0.0%	3.2%	1.7%	1.7%	2.0%	19.3%	14.4%
SE1	0.059	0.070	0.081	0.070	0.075	0.087	0.088
SE2	0.056	0.067	0.079	0.067	0.072	0.086	0.084
Lower percentile	0.050	0.037	0.016	0.041	0.029	0.047	0.034
Upper percentile	0.280	0.310	0.332	0.312	0.324	0.383	0.376
CP	0.957	0.961	0.954	0.953	0.950	0.961	0.960
Power	0.935	0.822	0.640	0.871	0.767	0.866	0.762

Note. M = censored on the mediator variable; X = censored on the input variable; Y = censored on the output variable; I = interval censored; L = left censored; CP = coverage probability; *RBE* = relative bias of estimate; *SE1* = mean of the standard error estimates of all the replications; *SE2* = standard deviation of the parameter estimates of all the replications. The 0% column contains the results from the data with 0% censoring.

With an even smaller sample size of 50, the mediation effects cannot be estimated well with 50% or more censored data on the output variable. The maximum RBE could be over 65% when the proportion of censored data was 75% on the output variable. From Tables 2 to 4, we can also find that the parameter estimates were relatively more accurate when the censoring was on the input variable or on the mediator variable than the output variable. Even with 75% of censored data on the input or mediator variable and a sample size of 50, the maximum RBE was only about 5.5%. In addition, regardless of the sample size and the proportion of censoring, the coverage probabilities were close to 95% for all the tested situations although the confidence interval lengths were much larger than the true confidence interval length with 75% censored data.

AN EMPIRICAL EXAMPLE

To illustrate the applications of the Bayesian Tobit mediation approach on estimating and testing

TABLE 4
Results from the Bayesian Tobit Approach on Estimating and Testing Mediation Effects with Censored Data ($n = 50$)

	0%	*M,I*	*M,L*	*X,I*	*X,L*	*Y,I*	*Y,L*
Proportion: 25%							
Estimate	0.152	0.153	0.154	0.152	0.153	0.160	0.159
RBE	0.1%	0.6%	1.2%	0.0%	0.5%	5.2%	4.3%
SE1	0.089	0.091	0.092	0.090	0.091	0.095	0.095
SE2	0.083	0.087	0.088	0.084	0.085	0.089	0.089
Lower percentile	0.004	0.000	−0.001	0.001	0.001	0.002	0.002
Upper percentile	0.348	0.356	0.360	0.354	0.357	0.375	0.372
CP	0.955	0.958	0.956	0.955	0.954	0.961	0.955
Power	0.523	0.496	0.482	0.508	0.504	0.509	0.500
Proportion: 50%							
Estimate	0.152	0.155	0.155	0.152	0.153	0.174	0.169
RBE	0.1%	2.2%	2.2%	0.2%	0.8%	14.4%	11.3%
SE1	0.089	0.097	0.101	0.094	0.098	0.111	0.109
SE2	0.083	0.094	0.097	0.088	0.091	0.104	0.105
Lower percentile	0.004	−0.008	−0.016	−0.003	−0.009	−0.003	−0.008
Upper percentile	0.348	0.372	0.381	0.366	0.374	0.431	0.420
CP	0.955	0.955	0.956	0.949	0.955	0.966	0.954
Power	0.523	0.438	0.379	0.472	0.433	0.465	0.436
Proportion: 75%							
Estimate	0.152	0.161	0.158	0.157	0.157	0.252	0.223
RBE	0.1%	5.5%	3.9%	3.2%	2.9%	65.4%	46.5%
SE1	0.089	0.107	0.126	0.107	0.119	0.204	0.189
SE2	0.083	0.104	0.123	0.101	0.110	0.270	0.204
Lower percentile	0.004	−0.021	−0.061	−0.013	−0.039	−0.019	−0.049
Upper percentile	0.348	0.399	0.437	0.405	0.429	0.758	0.681
CP	0.955	0.957	0.950	0.946	0.961	0.956	0.952
Power	0.523	0.357	0.225	0.413	0.290	0.439	0.343

Note. M = censored on the mediator variable; X = censored on the input variable; Y = censored on the output variable; I = interval censored; L = left censored; CP = coverage probability; *RBE* = relative bias of estimate; *SE1* = mean of the standard error estimates of all the replications; *SE2* = standard deviation of the parameter estimates of all the replications. The 0% column contains the results from the data with 0% censoring.

mediation effects with censored data, the model is applied to test whether verbal memory ability mediates the relationship between age and everyday functioning. The data used in this study are a subset of data from the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study (Willis et al., 2006). The ACTIVE study is a randomized controlled trial to examine the long-term outcomes of cognitive interventions on daily functioning of older individuals living independently. The multiple domain assessments were conducted across six waves on four groups including three treatment groups and a control group. In this study, only the data from the control group ($N = 698$) measured at the baseline occasion were used because of the demonstration purpose.

The participants' age ranged from 64 to 95 ($M = 74.04$, $SD = 6.05$). The verbal memory construct involved three measures of verbal memory ability: Hopkins Verbal Learning Test (Brandt, 1991), RAVLT (Rey, 1964), and Rivermead Behavioral Paragraph Recall Test (Wilson, Cockburn, & Baddeley, 1985). The verbal memory ability scores ($M = 30.59$, $SD = 6.22$) were generated from the averages of the three memory ability measures with equal weights for data reduction (Willis et al., 2006). The everyday functioning was measured by participant's self-rating of difficulty in completing cognitively demanding tasks involved in meal preparation, housework, finances, health maintenance, telephone use, and shopping (instrumental activities of daily living [IADL] difficulty). The values range from 0 to 20. In this sample, 54.58% of the participants ($M = 1.31$, $SD = 2.40$) reported 0, indicating that they did not have any difficulty in daily functioning. Compared to the participants who reported nonzero on IADL, we can consider that the participants who reported zero on IADL have negative difficulty in everyday life activity, or they can do more everyday life activities than the included activities in the scale. Therefore, the output variable, daily functioning measured by IADL difficulty, was substantially left censored and had floor effects (the proportion of censoring is 54.58%). No censoring was observed on the input (age) or the mediation (verbal memory ability) variables.

For the purpose of demonstration, both the Bayesian regular mediation analysis and the Bayesian Tobit mediation analysis were conducted with the empirical data. The results are provided in Table 5. To check convergence of the parameter estimates, the Geweke statistic value for each parameter was calculated utilizing the CODA function in R (Plummer, Best, Cowles, & Vines, 2006). The Geweke statistic values ranged between −1.96 and 1.96 for all the estimated parameters in the model from both approaches. Therefore, we can conclude that the results were likely to be convergent results.

Because there were no censored data on both the input and mediator variables, the estimates of a were identical from both models. However, the estimates of b and c' were very different. Especially, both the estimate and the standard error of ab using the Tobit approach were larger than the corresponding values from the regular model. This is consistent with the findings from the simulations with 50% of left censored data on the output variable in Tables 1 and 2. The posterior distributions of the product parameter (ab) from both models are displayed in Figure 2. From Figure 2, we can also see the differences in the results between two models. Based on the simulation results, the results from the Tobit analysis should be closer to the true values in this empirical example.

Based on the results from the Bayesian Tobit mediation analysis, we can conclude that the relationship between age and daily functioning among older adults is completely mediated by the verbal memory ability because the 95% confidence interval of the mediation effect does not include 0 and the direct effect, c', is not significantly different from 0.

TABLE 5
Empirical Results from Both Bayesian Regular and Bayesian Tobit Mediation Analysis ($n = 698$)

	Regular Analysis			*Tobit Analysis*		
Parameter	*Estimate*	*SE*	*Geweke*	*Estimate*	*SE*	*Geweke*
a	−0.434	0.038	0.543	−0.434	0.038	−0.497
b	−0.071	0.017	0.531	−0.116	0.034	0.561
ab	0.031	0.008	−0.537	0.050	0.015	−0.491
c'	0.021	0.017	0.771	0.051	0.033	0.754
σ^2_{eY}	5.497	0.296	0.559	17.710	1.571	−0.623
σ^2_{eM}	32.380	1.857	−0.491	32.320	1.844	−0.436
95% CI of ab		(0.016, 0.047)			(0.021, 0.082)	

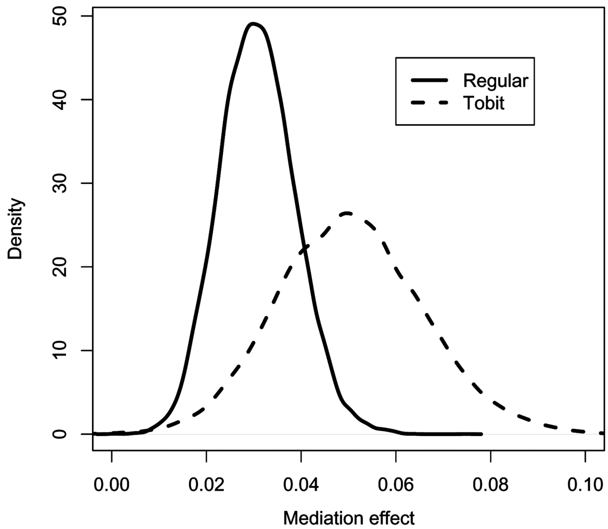

FIGURE 2 Posterior distributions of the product parameter from the Bayesian regular approach and the Bayesian Tobit approach.

DISCUSSION

This study first discussed the influences of censored data on mediation analysis. Through simulation, we found that censoring on any type of variable (input, mediation, or output) influenced both estimation and testing of mediation effects. Censoring on the output variable had greater influences than censoring on the other two variables. The simulation results also showed that both the point estimate and the standard error estimate of the mediation effect could be underestimated or overestimated and the coverage probability was lower than the desired level. With larger proportions of censored data, the biases were larger.

To deal with censored data in mediation analysis, the Tobit mediation model was introduced because it can be directly applied in estimating and testing mediation effects for censored data. The results from the Tobit mediation model showed that the point estimate of the mediation effect can be estimated very well with an adequate sample size (e.g., $n = 200$) even when the proportion of censored data was as high as 75%. The coverage probabilities of the estimated confidence intervals from all the tested simulations were close to the desired level regardless of the sample size, the type of censored variables, and the proportion of censored data. In terms of the standard error estimates, the results indicated that the standard error estimates with censored data were greater than the standard error estimates with 0% censored data, especially when the censoring proportion was relatively high, such as 50% or 75%. Furthermore, it is also shown that the power of testing mediation effects with censoring was lower than the power without censoring.

The simulation results have important implications, especially on when the Tobit mediation model can be used and when it might fail. From the results, we found that under certain conditions, the results from the Tobit approach can be biased, inefficient, or both. For example, when the sample size was 50 and the proportion of censored data on the output variable was either 50% or 75% or the sample size was 100 and the proportion of censored data on the output variable was 75%, the mediation effect cannot be very accurately recovered. Under these conditions, valid information from the observed data is very limited because of the small sample size and the high censoring proportion. We can compute the valid sample size by multiplying the total sample size by the proportion of uncensored data. Therefore, a rough guideline on applying the Tobit approach to simple mediation analysis is that the valid sample size should be at least 50 when the output variable is censored.

The Bayesian estimation method has been found to be a flexible method in estimating and testing mediation effect with censored data. First, it can obtain the empirical distribution of the mediation effect directly and thus can take into account the possible nonnormality of the mediation effect. Second, it is shown in the simulation that the Bayesian method can estimate model parameters and mediation effect accurately and precisely under certain conditions. Third, it can deal with censored data in more than one variable, including the input, output, and mediation variables relatively easily. In addition, the proposed procedure can be accordingly revised to conduct other path analysis with censored data relatively easily. Last but not least, the Bayesian estimation of the model can be implemented in available free software.

MacKinnon (2008) found that in some probit and logistic models, the difference $c - c'$ and product of coefficients ab approaches can lead to different results in testing mediation effects. It will be interesting to study whether the same problem occurs for Tobit mediation models in future studies.

In summary, censoring was found to play a nonignorable role in estimating and testing mediation effects. The proposed Bayesian Tobit mediation approach generally worked well to deal with the censoring problem in mediation analysis with enough valid information from the data. The findings from this study should encourage readers to use the proposed Bayesian Tobit mediation approach to estimate and test mediation effects with censored data.

REFERENCES

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Bollen, K. A., & Stine, R. A. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115–140.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist, 5*(2), 125–142.

Brown, R. L. (1992). Estimation problems with Type I censored response distributions in structural equation modeling. *Educational and Psychological Measurement, 52,* 325–336.

Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics, 51*, 79–99.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (1987). *California Verbal Learning Test: Adult version*. San Antonio, TX: Psychological Corporation.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7*(1), 1–26.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association, 82*(397), 171–185.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 721–741.

Geweke, J. (1992). Evaluating the accuracy of sampling–based approaches to the calculation of posterior moments. In J. O. B. J. M. Bernado, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics 4* (pp. 169–193). Oxford, UK: Clarendon.

Judd, C. M., & Kenny, D. A. (1981a). *Estimating the effects of social interventions.* New York: Cambridge University Press.

Judd, C. M., & Kenny, D. A. (1981b). Process analysis: Estimating mediation in evaluation research. *Evaluation Research, 5*, 602–619.

Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data* (2nd ed.). New York: Springer.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83–104.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*(1), 99–128.

Muthén, B. O. (1989). Tobit factor analysis. *British Journal of Mathematical and Statistical Psychology, 42*, 241–250.

Muthén, B. O. (1990). Moments of the censored and truncated normal distribution. *British Journal of Mathematical and Statistical Psychology, 43*, 131–143.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News, 6*(1), 7–11. Retrieved October 2007, from http://CRAN.R-project.org/doc/Rnews/

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717–731.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). San Francisco, CA: Jossey-Bass.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometria, 26*(1), 24–36.

Uttl, B. (2005). Measurement of individual differences: Lessons from memory assessment in research and clinical practice. *Psychological Science, 16*, 460–467.

Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research, 43*, 476–496.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology, 19*, 87–95.

Wechsler, D. (1987). *Wechsler Memory Scale—revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale—III.* San Antonio, TX: Psychological Corporation.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., et al. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA, 296*, 2805–2814.

Wilson, B., Cockburn, J., & Baddeley, A. (1985). *The Rivermead Behavioral Memory Test.* Reading, UK/Gaylord, MI: Thames Valley Test Co./National Rehabilitation Services.

Zhang, Z., McArdle, J. J., Wang, L., & Hamagami, F. (2008). A SAS interface for Bayesian analysis with WinBUGS. *Structural Equation Modeling, 15*, 705–728.

Zhang, Z., & Wang, L. (2008). Methods for evaluating mediation effects: Rationale and comparison. In K. Shigemasu, A. Okada, T. Imaizumi, & T. Hoshino (Eds.), *New trends in psychometrics* (pp. 595–604). Tokyo: Universal Academy Press.

APPENDIX A
THE PRIOR AND CONDITIONAL POSTERIOR DISTRIBUTIONS

The prior distributions are:

$$\mu_X \sim N(\mu_{X0}, \sigma^2_{X0}); \sigma^2_X \sim IG(k_{X0}/2, s_{X0}/2)$$

$$i_1 \sim N(i_{10}, \sigma^2_{i10}); b \sim N(b_0, \sigma^2_{b0}); c' \sim N(c_0, \sigma^2_{c0}); \sigma^2_{eY} \sim IG(k_{Y0}/2, s_{Y0}/2)$$

$$i_2 \sim N(i_{20}, \sigma^2_{i20}); a \sim N(a_0, \sigma^2_{a0}); \sigma^2_{eM} \sim IG(k_{M0}/2, s_{M0}/2)$$

Therefore, the conditional posterior distributions of the parameters in the Tobit mediation model are:

$$x^*_{c,i}|\mu_X, \sigma^2_X, a, b, c', i_1, i_2, \sigma^2_{eM}, \sigma^2_{eY}, y^*_i, m^*_i \sim$$

$$NX\left(\frac{\frac{\mu_X}{\sigma^2_X} + \frac{(m^*_i - i_2)a}{\sigma^2_{eM}} + \frac{(y^*_i - i_1 - bm^*_i)c'}{\sigma^2_{eY}}}{\frac{1}{\sigma^2_X} + \frac{a^2}{\sigma^2_{eM}} + \frac{c'^2}{\sigma^2_{eY}}}, \frac{1}{\frac{1}{\sigma^2_X} + \frac{a^2}{\sigma^2_{eM}} + \frac{c'^2}{\sigma^2_{eY}}}\right)$$

$$m^*_{c,i}|a, b, c', i_1, i_2, \sigma^2_{eM}, \sigma^2_{eY}, y^*_i, x^*_i \sim NM\left(\frac{\frac{ax^*_i + i_2}{\sigma^2_{eM}} + \frac{(y^*_i - i_1 - c'x^*_i)b}{\sigma^2_{eY}}}{\frac{1}{\sigma^2_{eM}} + \frac{b^2}{\sigma^2_{eY}}}, \frac{1}{\frac{1}{\sigma^2_{eM}} + \frac{b^2}{\sigma^2_{eY}}}\right)$$

$$y^*_{c,i}|i_1, b, c', \sigma^2_{eY}, x^*_i, m^*_i \sim NY\left(i_1 + bm^*_i + c'x^*_i, \sigma^2_{eY}\right)$$

$$\mu_X|\sigma^2_X, x^* \sim N\left(\frac{\frac{\sum x^*_i}{\sigma^2_X} + \frac{\mu_{X0}}{\sigma^2_{X0}}}{\frac{n}{\sigma^2_X} + \frac{1}{\sigma^2_{X0}}}, \frac{1}{n/\sigma^2_X} + \frac{1}{\sigma^2_{X0}}\right)$$

$$\sigma^2_X|\mu_X, x^* \sim IG\left(\frac{n + k_{X0}}{2}, \frac{\sum(x^*_i - \mu_X)^2 + s_{X0}}{2}\right)$$

$$a|x^*, m^*, i_2, \sigma^2_{eM} \sim N\left(\frac{\frac{\sum(m^*_i - i_2)x^*_i}{\sigma^2_{eM}} + \frac{a_0}{\sigma^2_{a0}}}{\frac{\sum x^{*2}_i}{\sigma^2_{eM}} + \frac{1}{2\sigma^2_{a0}}}, \frac{1}{\frac{\sum x^{*2}_i}{\sigma^2_{eM}} + \frac{1}{2\sigma^2_{a0}}}\right)$$

$$b|x^*, m^*, y^*, i_1, c', \sigma^2_{eY} \sim N\left(\frac{\frac{\sum(y^*_i - c'x^*_i - i_1)m^*_i}{\sigma^2_{eY}} + \frac{b_0}{\sigma^2_{b0}}}{\frac{\sum m^{*2}_i}{\sigma^2_{eY}} + \frac{1}{\sigma^2_{b0}}}, \frac{1}{\frac{\sum m^{*2}_i}{\sigma^2_{eY}} + \frac{1}{\sigma^2_{b0}}}\right)$$

$$c'|x^*, m^*, y^*, i_1, b, \sigma^2_{eY} \sim N\left(\frac{\frac{\sum(y^*_i - bm^*_i - i_1)x^*_i}{\sigma^2_{eY}} + \frac{c_0}{\sigma^2_{c0}}}{\frac{\sum x^{*2}_i}{\sigma^2_{eY}} + \frac{1}{\sigma^2_{c0}}}, \frac{1}{\frac{\sum x^{*2}_i}{\sigma^2_{eY}} + \frac{1}{\sigma^2_{c0}}}\right)$$

$$\sigma^2_{eM}|x^*, m^*, a, i_2 \sim IG\left(\frac{n + k_{Mo}}{2}, \frac{\sum(m^*_i - i_2 - ax^*_i)^2 + s_{Mo}}{2}\right)$$

$$\sigma^2_{eY}|x^*, y^*, m^*, c', b, i_1 \sim IG\left(\frac{n + k_{Y0}}{2}, \frac{\sum(y^*_i - i_1 - c'x^*_i - bm^*_i)^2 + s_{Y0}}{2}\right)$$

APPENDIX B
WinBUGS CODES FOR REGULAR AND TOBIT MEDIATION MODELS

Regular Mediation Analysis

```
model{
for (i in 1: 200 ) {
m[i]~ dnorm(mum[i], Inv_sig2_em)
mum[i]<-im+a*x[i]
y[i]~dnorm(muy[i], Inv_sig2_ey)
muy[i]<-iy+c*x[i]+b*m[i]
}
#priors
a~dnorm(0, 1.0E-6)
b~dnorm(0, 1.0E-6)
c~dnorm(0, 1.0E-6)
im~dnorm(0, 1.0E-6)
iy~dnorm(0, 1.0E-6)
Inv_sig2_ey~dgamma(1.0E-3, 1.0E-3)
Inv_sig2_em~dgamma(1.0E-3, 1.0E-3)
#parameter transformation
Sig2_ey<-1/Inv_sig2_ey
Sig2_em<-1/Inv_sig2_em
product<-a*b
}
```

Tobit Mediation Analysis

```
model{
for (i in 1: 200 ) {
x[i]~ dnorm(mux,taux)I(lowerx[i],upperx[i])
mum[i]<-im+a*x[i]
m[i]~ dnorm(mum[i],taum)I(lowerm[i],upperm[i])
muy[i]<- iy+c*x[i]+b*m[i]
y[i]~ dnorm(muy[i],tauy)I(lowery[i],uppery[i])}
#priors
a~dnorm(0, 1.0E-6)
b~dnorm(0, 1.0E-6)
c~dnorm(0, 1.0E-6)
interceptm~dnorm(0,1.0E-6)
intercepty~dnorm(0,1.0E-6)
mux~dnorm(0,1.0E-6)
taum~dgamma(0.001,0.001)
tauy~dgamma(0.001,0.001)
taux~dgamma(0.001,0.001)
#parameter transformation
varm<-1/taum
vary<-1/tauy
product<-a*b
}
```