Treatment Effects and Informative Missingness with an Application to Bank Recapitalization Programs

By ANGELA VOSSMEYER*

This paper extends conventional treatment models to allow for missing outcomes, endogeneity and discrete data. The new methodology is employed to evaluate the effectiveness of bank recapitalization programs and their ability to resuscitate the financial system.

Many policies and programs that are the foci of economic research operate under selection mechanisms or decision structures that must be accommodated by econometric models to ensure proper inference. However, the complexities involved in modeling selection equations and treatment response data have restricted attention to single equation models. The dangers of improper modeling are bias and misrepresentation of the population of interest. These issues are prevalent in the analysis of lender of last resort (LOLR) policies, which underlies the application in this paper.

Treatment models consider two subgroups in the data, the treated group and the control, or untreated, group. This formulation, when applied to an LOLR study, divides the sample into banks that receive loans from the LOLR and banks that do not receive loans. Complications arise because an initial selection mechanism, the application step of the recapitalization process, is ignored. Avoiding the question of whether the bank applied for assistance from the LOLR would erroneously group banks that do not apply for assistance with those that are declined assistance. Thus, the untreated group comprises the most and least healthy banks leading to a fundamental misspecification. This paper remedies the problem by developing and implementing a multivariate treatment effect model for non-randomly selected data to offer a more complete framework for evaluating the impact of LOLR programs.

The existing literature on LOLR policies and bank recapitalization is mixed. Research in favor of recapitalization programs finds that LOLR policies play a positive role in reducing bank failures and improving monetary conditions (Butkiewicz, 1995; Richardson and Troost, 2009). Loose lending policies can prevent the spread of contagion, bank runs, and mass liquidation. Other studies find that LOLR policies can be harmful either by restricting banks' good collateral or by creating moral hazard incentives for banks to take on excessive risk (Mason, 2001; Mishkin, 2006). These issues have been deliberated since the concept of LOLR was described by Bagehot (1873). Bagehot states that monetary authorities, in the face of panic, should lend unsparingly at a penalty rate to illiquid but solvent banks. This mechanism should prevent struggling healthy banks from falling victim to undue deposit losses, bank runs, and insolvency.

This paper focuses on the Reconstruction Finance Corporation (RFC) as the LOLR during the Great Depression. The RFC started in 1932 and became one of the largest LOLR programs ever implemented. The empirical analysis of these programs poses a number of challenges because regulator data is generally not publicly available, and modeling involves a difficult decision structure which is at the intersection of treatment effect models, sample selection, endogeneity, and discrete data modeling. This paper adds to the existing literature by employing a novel bank-level data set and developing a new methodology to jointly model a bank's decision to apply for a loan from the LOLR, the LOLR's decision to approve the loan, and the bank's success a few years after the disbursements.

I. Methodology

Although models for sample selection and treatment effects are used frequently on their own, techniques that incorporate and jointly model both are lacking in the literature. To address this deficiency, this paper employs a

* Department of Economics, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100; e-mail: angela.vossmeyer@uci.edu. Special thanks to: Ivan Jeliazkov, Gary Richardson, David Brownstone, and Dale Poirier.

Bayesian framework for treatment effect modeling while dealing with the missing data that occur from sample selection, shown in Figure 1.

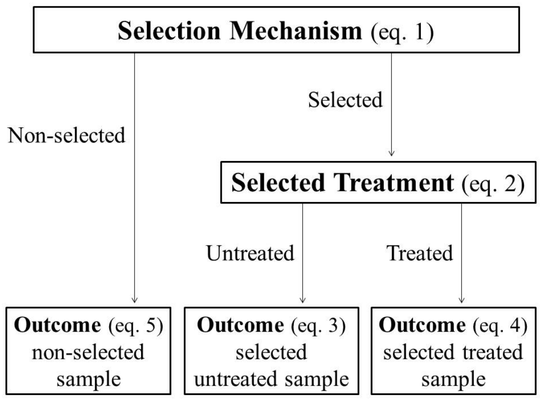

Figure 1: Treatment model with sample selection

Figure 1 presents the multi-step data generation process leading to the observed data. This framework can be recognized as the decision structure employed in many programs consisting of an application step and an approval step. The initial selection mechanism is observed for the entire sample. The selected sample then enters a selected treatment stage followed by a set of potential outcomes or treatment responses. The model considered here differs from the existing literature by acknowledging whether a bank opts into or out of treatment to disentangle the information content in not selecting. Modeling this additional selection mechanism offers revealing features of the data.

The model stemming from Figure 1 contains 5 equations of interest – 1 selection mechanism, 1 selected treatment and 3 treatment response outcomes for the different subsets of the sample: the non-selected sample, the selected untreated sample, and the selected treated sample. For the banking application, the non-selected sample comprises banks that do not apply for assistance from the LOLR, the selected untreated includes banks that apply and are denied assistance, and the selected treated comprises banks that apply and are granted assistance. In detail, the equations for banks $i = 1, \ldots, n$ are given by:

SELECTION MECHANISM (*Decision to apply for a loan*)– always observed:

$$(1) \qquad y^*_{i1} = \mathbf{x}'_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1}$$

SELECTED TREATMENT (*LOLR's decision to approve the loan*)– observed for the selected sample, missing for the non-selected sample:

$$(2) \qquad y^*_{i2} = \mathbf{x}'_{i2}\boldsymbol{\beta}_2 + \varepsilon_{i2}$$

TREATMENT RESPONSES (*Bank failure or success*)– only 1 equation is observed:

$$\begin{aligned}
(3)\ y^*_{i3} &= (\mathbf{x}'_{i3}\ y_{i1})\boldsymbol{\beta}_3 + \varepsilon_{i3}, \text{ (selected untreated)}\\
(4)\ y^*_{i4} &= (\mathbf{x}'_{i4}\ y_{i1}\ y_{i2})\boldsymbol{\beta}_4 + \varepsilon_{i4}, \text{ (selected treated)}\\
(5)\ y^*_{i5} &= \mathbf{x}'_{i5}\boldsymbol{\beta}_5 + \varepsilon_{i5}, \text{ (non-selected)}
\end{aligned}$$

The model is characterized by 5 dependent variables of interest where $\mathbf{y}^*_i$ are the continuous latent data and $\mathbf{y}_i$ are the corresponding discrete observed data. In the application, the latent variables y^*_{ij} relate to the observed censored outcomes y_{ij} by $y_{ij} = y^*_{ij} I(y^*_{ij} > 0)$ for equations $j = 1, \ldots, 5$ (Tobin, 1958). This general system, however, can take outcome variables that are continuous, binary, censored or ordered.

Data missingness restricts the model to systems of 2 or 3 equations, depending on the subsample to which the observation belongs and highlights the presence of non-identified parameters that will be examined shortly. For observations in the non-selected sample (indexed by N_1), $\mathbf{y}_{ic} = (y_{i1}, y_{i5})'$ is the vector of outcomes for the observed system of equations and (y_{i2}, y_{i3}, y_{i4}) are not observed. Learning from observations in the selected untreated sample (indexed by N_2) coincides with the system of equations for which $\mathbf{y}_{id} = (y_{i1}, y_{i2}, y_{i3})'$ is observed, and (y_{i4}, y_{i5}) are not observed . The observed system of equations for the selected treated sample (indexed by N_3) contains the outcomes $\mathbf{y}_{ia} = (y_{i1}, y_{i2}, y_{i4})'$ where y_{i3} and y_{i5} are not observed.

The exogenous covariates $\mathbf{x}_i$=($\mathbf{x}_{i1}$, $\mathbf{x}_{i2}$, $\mathbf{x}_{i3}$, $\mathbf{x}_{i4}$, $\mathbf{x}_{i5}$) are needed only when their corresponding equations are observed. The model assumes that the errors ε_{ij} for $j = 1, \ldots, 5$ have a multivariate normal distribution $\mathcal{N}_5(0, \boldsymbol{\Omega})$ where,

$$\boldsymbol{\Omega} = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} & \Omega_{24} & \cdot \\ \Omega_{31} & \Omega_{32} & \Omega_{33} & \cdot & \cdot \\ \Omega_{41} & \Omega_{42} & \cdot & \Omega_{44} & \cdot \\ \Omega_{51} & \cdot & \cdot & \cdot & \Omega_{55} \end{pmatrix}.$$

Note that there are 11 unique elements in $\boldsymbol{\Omega}$ that can be estimated, whereas the remaining ones are non-identified parameters due to the missing outcomes.

Following Chib, Greenberg and Jeliazkov (2009) and Chib (2007), the likelihood is defined in terms of the 3 subsets of the sample without computations for the unobserved components of the model. Let $\boldsymbol{\mu}_j$ define the mean in equations $j = 1, \ldots, 5$. The aforementioned subscripts c, d and a identify the observed system of equations so, $vech(\boldsymbol{\Omega}_c) = (\Omega_{11}, \Omega_{15}, \Omega_{55})'$, $\boldsymbol{\mu}_c = (\mu_1, \mu_5)'$, $vech(\boldsymbol{\Omega}_d) = (\Omega_{11}, \Omega_{21}, \Omega_{22}, \Omega_{31}, \Omega_{32}, \Omega_{33})'$, $\boldsymbol{\mu}_d = (\mu_1, \mu_2, \mu_3)'$, $vech(\boldsymbol{\Omega}_a)$=($\Omega_{11}$,$\Omega_{21}$,$\Omega_{22}$, Ω_{41},Ω_{42},Ω_{44})', and $\boldsymbol{\mu}_a = (\mu_1, \mu_2, \mu_4)'$. [1] This notation can be used in constructing the likelihood $f(\mathbf{y}|\boldsymbol{\theta}) = \int f(\mathbf{y}, \mathbf{y}^*|\boldsymbol{\theta}) d\mathbf{y}^*$ where $\boldsymbol{\theta}$ is all model parameters, and $f(\mathbf{y}, \mathbf{y}^*|\boldsymbol{\theta})$ is given by

$$\prod_{i \in N_1} f_N(\mathbf{y}^*_{ic}|\boldsymbol{\mu}_c, \boldsymbol{\Omega}_c) \times \prod_{i \in N_2} f_N(\mathbf{y}^*_{id}|\boldsymbol{\mu}_d, \boldsymbol{\Omega}_d) \times \prod_{i \in N_3} f_N(\mathbf{y}^*_{ia}|\boldsymbol{\mu}_a, \boldsymbol{\Omega}_a).$$

The discreteness of multiple outcome variables render this likelihood analytically intractable and hence estimation relies on simulation-based techniques. Standard semi-conjugate priors are applied where $\boldsymbol{\beta}$ has a joint normal distribution and (independently) $\boldsymbol{\Omega}$ has an inverted Wishart distribution. Combining the likelihood and priors leads to a posterior distribution which is simulated by Markov chain Monte Carlo (MCMC) methods. For computational efficiency, a collapsed Gibbs sampler with data augmentation is employed which follows from Chib, Greenberg and Jeliazkov (2009) and Li (2011). A complete description of the estimation algorithm is offered in the online appendix. [2]

II. Data

This paper employs two novel bank-level data sets: RFC data and bank balance sheet data. The RFC data set is collected from the "RFC Card Index of Loans Made to Banks and Railroads, 1932-1957" which was acquired from the National Archives in College Park, Maryland. The cards report the name of the borrower, request and amount of loan, and whether the loan was approved or declined. Further information on each loan is obtained from the "Paid Loan Files" and "Declined Loan Files" which include the exact information the regulators had on each bank and the original examiner's report on each decision. This data set is merged with a separate data set constructed from the *Rand McNally Banker's Directory* which describes balance sheets, correspondent relationships and characteristics for all banks in a given state. This information identifies the non-selected, or non-applicant sample. Additional data are gathered from the 1930 U.S. census of agriculture, manufacturing and population which describe the characteristics of the county and a bank's business environment.

The data are applied to the 5 equation model as follows: the outcome variable for equation 1 is total amount of RFC assistance requested by December 1933. This outcome is censored with point mass at zero for banks that did not apply for assistance and a continuous distribution for the different loan amounts requested. The outcome variable for equation 2 is the total amount of RFC assistance approved. This outcome is also censored with point mass at zero for banks that were declined assistance and a continuous distribution for the approved loan amounts.

The outcome for equations 3–5 is the amount of "loans and discounts" (hereafter, LD) for each bank taken from its January 1935 balance sheet. The year 1935 is selected because the intervening years allowed banks to use their relief funds. The outcome for each equation is censored with point mass at zero for banks that failed since the time of the loan applications and a continuous distribution with LD representing the bank's health and the state of the local economy. 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 and Mason, 2003).

The sample includes all banks operating in 1932 in Alabama, Arkansas, Mississippi, Michigan and Tennessee. The sample consists of 1,794 banks, of which 908 banks applied for RFC assistance, and 800 of those were approved while 108 were declined assistance. Covariates include financial ratios, charters, memberships, departments, correspondent relationships, market shares, and county characteristics.

[1] $vech$ extracts the unique elements of a symmetric matrix.
[2] http://sites.uci.edu/vossmeyer.

III. Results

Analysis of the resulting parameter estimates from the multivariate treatment model is complicated by the discreteness of the outcome variables. Interpretation is afforded with covariate and treatment effect calculations which are important for understanding the model and for determining the impact of a change in a covariate.

The coefficient on the endogenous covariate y_{i2} in equation 4 (hereafter, β_{RFC}) is the key estimate of interest. After controlling for a bank's health, environment and contagion channels, this covariate reflects the impact of RFC assistance on bank lending. The basic result indicates this parameter is positive with a credibility interval that does not include zero. To calculate how a change in RFC lending transfers to bank lending, the marginal effect is averaged over both observations and MCMC draws. The marginal effect for β_{RFC} is 0.571 which can be interpreted as, \$10,000 of RFC assistance translates to \$5,710 of LD in 1935. This result accords well with the deposit-to-loan ratios during the 1930s and during banking panics, in general. The money that was not converted to loans was likely kept in cash reserves to prepare for a bank run. RFC assistance was effectively pushed beyond banks trickling into local economies, thus restoring confidence in the financial system.

To deduce treatment effects, 2 scenarios are considered. The first case is the difference in the probability of bank failure if the RFC did not offer any assistance. To see how removing the treatment from the treated banks affects bank success, two probabilities need to be computed, $\Pr(y_{i4} = 0|x_{i4}, y_{i2}^{\ddagger}, \boldsymbol{\theta})$ and $\Pr(y_{i4} = 0|x_{i4}, y_{i2}^{\dagger}, \boldsymbol{\theta})$ where $y_{i2}^{\ddagger}$ represents zero RFC assistance, and $y_{i2}^{\dagger}$ represents the original treatment. Thus, interest lies in how a bank's probability of failure changes if the RFC never approved any loans. Formally, the objective is to obtain a sample of draws and evaluate

$$\{\Pr(y_{i4} = 0|y_{i2}^{\ddagger}) - \Pr(y_{i4} = 0|y_{i2}^{\dagger})\}$$

$$= \int \{\Pr(y_{i4} = 0|\mathbf{x}_{i4}, y_{i2}^{\ddagger}, \boldsymbol{\theta}) -$$

$$\Pr(y_{i4} = 0|\mathbf{x}_{i4}, y_{i2}^{\dagger}, \boldsymbol{\theta})\}\pi(\mathbf{x}_{i4})\pi(\boldsymbol{\theta}|\mathbf{y})d\mathbf{x}_{i4}d\boldsymbol{\theta}.$$

The result gives the expected difference in the computed pointwise probabilities as $y_{i2}^{\dagger}$ is changed to $y_{i2}^{\ddagger}$ (Jeliazkov, Graves and Kutzbach, 2008). Computation of these probabilities is afforded by employing the CRT method, developed in Jeliazkov and Lee (2010). The results indicate that the probability difference equals 0.126. In other words, if the RFC did not offer any assistance, the probability of bank failure for the selected treated sample (approved banks) increases by 12.6 percentage points.

The second treatment effect to consider is how RFC assistance could have changed the outcomes for banks that were declined loans. For this scenario, the RFC approved loans are equated to the amounts requested on declined banks' applications. Two probabilities to consider are, $\Pr(y_{i3} = 0|x_{i3}, y_{i2}^{\ddagger}, \boldsymbol{\theta})$ and $\Pr(y_{i3} = 0|x_{i3}, y_{i2}^{\dagger}, \boldsymbol{\theta})$ where $y_{i2}^{\ddagger}$ represents declined loans (the original case), and $y_{i2}^{\dagger}$ represents the scenario where the RFC approved the full requested amounts. This situation displays the difference in the probability of failure if the RFC approved applications for the selected untreated sample (declined banks). The results show, $\{\Pr(y_{i3} = 0|y_{i2}^{\ddagger}) - \Pr(y_{i3} = 0|y_{i2}^{\dagger})\} = 0.025$. If the RFC assisted banks that were declined loans, the probability of failure for the selected untreated sample decreases by 2.5 percentage points. RFC loans are almost 5 times more effective in the approved bank subsample. The banks the RFC declined to assist were helpless because full assistance from the RFC would not have had a major impact on their ability to survive and thrive in the economy.

The results of the two scenarios are clear. LOLR policies and bank recapitalization aided a bank's survival if the bank was healthy enough to receive a loan. Once non-randomly appointed to the treated group, banks that received RFC loans converted their relief funds to LD supporting local economies. The results also indicate that the selection procedures adopted by the RFC were successful. Assistance to all struggling banks would have been wasteful because most of the untreated banks were not healthy enough to have benefitted from an influx of funds. Therefore, proper consideration of the decision structure and composition of the treatment and control groups are of fundamental importance to evaluating the effectiveness of LOLR programs.

IV. Concluding Remarks

This paper presents a methodological framework for multivariate treatment effect models in the presence of sample selection and discrete data. The model is applicable to a multitude of problems prevalent in economics including modeling the effectiveness of job training and housing programs, health treatments, education policies, credit approval decisions, and many others. On the technical side, the methodology developed here is computationally efficient, and has low storage costs.

The methods established in this paper are applied to the analysis of LOLR regulation. The results indicate that bank recapitalization is effective at decreasing the probability of bank failure and stimulating bank lending. The use of the multivariate treatment effect model is extremely important to the findings because the results vary for the different subgroups of banks and selection into these groups is non-random. Although RFC assistance was beneficial for the treated group, it would have been minimally helpful for banks that were declined assistance because their economic condition was too severe.

Studying the RFC is an important and relevant topic because the RFC was used as a model for the current program, the Troubled Asset Relief Program (TARP). Further research on LOLR policies should focus on the multi-step decision mechanisms that place banks into different policies and programs to answer questions including whether and to what extent these programs stabilize the economy or simply privatize the gains and nationalize the loses. Overall, this model offers practical estimation tools to unveil new answers to questions involving sample selection and treatment response data.

REFERENCES

Bagehot, Walter. 1873. *Lombard Street: A Description of the Money Market.* London, NY: H.S. King.

Bernanke, Ben S. 1983. "Nonmonetary Effects of the Financial Crisis in Propagation of the Great Depression." *American Economic Review*, 73(3): 257–276.

Butkiewicz, James. 1995. "The Impact of Lender of Last Resort during the Great Depression: The Case of the Reconstruction Finance Corporation." *Explorations in Economic History*, 32(2): 197–216.

Calomiris, Charles, and Joseph Mason. 2003. "Consequences of Bank Distress During the Great Depression." *American Economic Review*, 93(3): 937–947.

Chib, Siddhartha. 2007. "Analysis of Treatment Response Data without the Joint Distribution of Potential Outcomes." *Journal of Econometrics*, 140: 401–412.

Chib, Siddhartha, Edward Greenberg, and Ivan Jeliazkov. 2009. "Estimation of Semiparametric Models in the Presence of Endogeneity and Sample Selection." *Journal of Computational and Graphical Statistics*, 18: 321–348.

Jeliazkov, Ivan, and Esther Hee Lee. 2010. "MCMC Perspectives on Simulated Likelihood Estimation." *Advances in Econometrics*, 26: 3–39.

Jeliazkov, Ivan, Jennifer Graves, and Mark Kutzbach. 2008. "Fitting and Comparison of Models for Multivariate Ordinal Outcomes." *Advances in Econometrics*, 23: 115–156.

Li, Phillip. 2011. "Estimation of sample selection models with two selection mechanisms." *Computational Statistics and Data Analysis*, 55(2): 1099–1108.

Mason, Joseph R. 2001. "Do Lender of Last Resort Policies Matter? The Effects of Reconstruction Finance Corporation Assistance to Banks During the Great Depression." *Journal of Financial Services Research*, 20(1): 77–95.

Mishkin, F. 2006. "How Big a Problem is Too Big to Fail? A Review of Gary Stern and Ron Feldman's Too Big to Fail: The Hazards of Bank Bailouts." *Journal of Economic Literature*, 44(4): 988–1004.

Richardson, Gary, and William Troost. 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(6): 1031–1073.

Tobin, James. 1958. "Estimation of relationships for limited dependent variables." *Econometrica*, 26(1): 24–36.