ARE ALL U.S. CREDIT UNIONS ALIKE?

A GENERALIZED MODEL OF HETEROGENEOUS TECHNOLOGIES WITH ENDOGENOUS SWITCHING AND FIXED EFFECTS

Emir Malikov†, Diego A. Restrepo-Tobón‡, Subal C. Kumbhakar§

August 15, 2013

Abstract

This paper formally models heterogeneity in credit unions' production technologies as evidenced by (endogenously selected) differing output mixes. We show that failure to account for this observed heterogeneity is likely to lead to biased, inconsistent estimates of credit unions' technology and potentially misleading results about the industry structure. The estimates are also likely to be biased if one overlooks unobserved credit union specific effects, as customarily done in the literature. To address these concerns, we develop a generalized model of endogenous switching with polychotomous choice that is able to accommodate fixed effects in both the technology selection and the outcome equations. We use this model to estimate returns to scale for the U.S. retail credit unions from 1994 to 2011. Unlike recent studies, we find that not all credit unions enjoy increasing returns to scale. A non-negligible number of large institutions operate at decreasing returns to scale, indicating that they should either cut back in size or switch to a different technology by adjusting the output mix.

Keywords: Credit Union, Fixed Effects, Panel Data, Returns to Scale, Selection, Switching Regression

JEL Classification: C33, C34, G21

† Department of Economics, State University of New York at Binghamton, Binghamton, NY; *Email*: emalikov@binghamton.edu.

‡ Department of Finance, EAFIT University, Medellín, Colombia; *Email*: drestr16@eafit.edu.co.

§ Department of Economics, State University of New York at Binghamton, Binghamton, NY; *Email*: kkar@binghamton.edu.

We are thankful to Jeffrey Wooldridge and Alfonso Flores-Lagunes for helpful comments and suggestions. Restrepo-Tobón acknowledges financial support from the Colombian Administrative Department of Sciences, Technology and Innovation; the Colombian Fulbright Commission; and EAFIT University.

'All happy families are alike; every unhappy family is unhappy in its own way.'
Leo Tolstoy, *Anna Karenina*.

1 INTRODUCTION

U.S. credit unions continue to prosper despite the decline in their relative advantages over commercial banks. Factors such as increasing availability of credit information from national credit-reporting bureaus, establishment of the federal deposit insurance fund for credit unions and the growth in credit card lending by larger financial institutions have significantly eroded conventional benefits of doing business at the local, small-scale level (Petersen & Rajan, 2002; Walter, 2006; Wheelock & Wilson, 2011). This has motivated credit unions to evolve. With the authorization to issue long-term mortgage loans in 1977 and the passage of the Credit Union Membership Access Act of 1998 which empowered them to widen and diversify their membership scope, credit unions have grown significantly in an attempt to compensate for the loss of traditional competitive advantages by capitalizing on economies of scale. The industry has been undergoing a wave of mergers and acquisitions: within the past decade alone there was a total of 2,464 mergers (2002-2011). Over these ten years, the average size of (federally-insured) credit unions has increased from $57.5 million to $135.8 million in assets. As of the end of 2011, the industry accounted for about a trillion dollars in assets and more than 92 million members (authors' calculations based on NCUA, 2011).

Several studies have investigated the performance of U.S. credit unions.[1] However, to our knowledge, no attempt has been made to formally model credit unions' technologies taking into consideration their differing output mixes. This limits our understanding of the industry structure, its evolution and the potential impact of alternative policies. All existing studies have encountered the same problem, namely, the presence of a large number of observations for which the reported values of outputs are zeros. This issue has been handled either by linearly aggregating different types of outputs into larger bundles (Fried, Lovell & Yaisawarng, 1999; Frame & Coelli, 2001; Wheelock & Wilson, 2011, 2013) or by replacing zero outputs with an arbitrarily chosen small positive number (Frame *et al.*, 2003). There may however be concerns whether these methods are appropriate since they do not recognize that the existence of zero-value outputs provides valuable information regarding the choice of the production technology by credit unions.

[Insert Table 1 here]

To preview the importance of modeling the choice of credit unions' technology properly (which we discuss in detail in Section 2), consider Table 1 which presents the number of retail credit unions in each year between 1994 and 2011 with zero values reported for some (or all) of the four outputs commonly considered in the literature. All credit unions[2] report non-zero values for consumer loans (*y3*) which historically have been the main product of credit unions. However,

[1] See Wheelock and Wilson (2011, 2013) and the references therein.
[2] With the exception of a single entity.

there is a strikingly large number of credit unions that offer no real estate ($y1$) or business loans ($y2$) to their members throughout the years we consider. This evidence favors our view that not all credit unions are alike. Given that the output mix differs across units and over time, a substantial time-persistent heterogeneity might exist among credit unions.

We view this observed heterogeneity as an outcome of an endogenous choice made by credit unions. They decide what range of services to offer to their members and choose the appropriate technology to provide them. Thus, it is likely that the production technology which a credit union employs varies with its output mix. To our knowledge, this technological heterogeneity (defined by the output mix) has been either assumed to be exogenous and/or completely taken for granted in all previous studies of the credit union technologies. The aggregation of outputs into broader categories to solve the zero-output problem, so often practiced in the literature, constitutes the loss of information in both econometric and economic senses. The results previously reported in the literature are therefore likely to be misleading since the used econometric models ignore the time-persistent heterogeneity arising from the endogenous selection of credit unions' technologies.[3]

Heterogeneity among credit unions is unlikely to be limited to the technology they use only; each credit union is unique in its operations. Ignoring this unobserved heterogeneity when estimating credit unions' technology (which is customary in the existing literature[4]) may produce inconsistent estimates since unobserved heterogeneity is likely to be correlated with covariates in the estimated equation. While such credit-union-specific unobserved effects cannot be accounted for in a cross-sectional setting due to the incidental parameters problem, we address this issue in our case by taking advantage of the panel structure of the U.S credit union data.

In this paper, we address the above concerns by developing a unified framework that allows estimation of credit union technologies that is robust to (i) misspecification due to an *a priori* assumption of homogenous technology, (ii) selectivity bias due to ignoring the endogeneity in technology selection, and (iii) endogeneity (omitted variable) bias due to a failure to account for unobserved union-specific effects that are correlated with covariates in the estimated equations.

---

[3] We acknowledge that the issue of heterogeneity among credit unions has been also addressed (although from a somewhat different perspective) in Wheelock and Wilson (2011) who estimate credit unions' cost function via kernel methods, thus avoiding any functional specification for the underlying technology and obtaining observation-specific estimates of the cost function. A kernel regression indeed permits credit unions' technologies to be completely heterogeneous with respect to covariates included in the regression. However, the aggregation of all types of loans into a single output, which Wheelock and Wilson (2011) resort to, does not allow them to account for the heterogeneity resulting from differing output mixes which this paper emphasizes. The authors do include two indicator variables in their regressions to control for zero-value (disaggregated) outputs. While the latter partly resolves the issue, the information on output-type-specific variation is still being lost which is likely to affect the results on scale economies reported in the paper. More importantly, similar to the rest of the literature, Wheelock and Wilson (2011) do not consider a likely possibility of differing output mixes being endogenously determined by credit unions which, as we show in this paper among other things, may result in severely distorted results due to the unaddressed selectivity bias. That is, the above-mentioned indicator variables used by Wheelock and Wilson (2011) are likely to be endogenous.

[4] To our knowledge, Frame *et al*. (2003) is the only study which attempts to estimate (homogenous) credit unions' technology using panel data while allowing for unobserved heterogeneity among unions. However, the latter is modelled as random effects under a strong assumption of its exogeneity which is unlike to hold in practice.

However, the estimation of such a model is not trivial. As we demonstrate in Section 2, the data indicate that 99% of all U.S. retail credit unions employ one of the three technologies associated with different output mixes offered by these institutions from 1994 to 2011. This suggests that a credit union's choice of technology is of a polychotomous nature. In the cross-sectional setting, an econometric strategy appropriate for such a problem would be a switching regression model [Maddala's (1983) terminology] under polychotomous choice considered by Lee (1978, 1982, 1983 and 1995). The latter method however does not have the ability to take into account unobserved effects. At the same time, the existing literature on panel data selection models with unobserved heterogeneity focuses mainly on binary selection, and few papers allow for fixed-effect type heterogeneity in *both* the outcome and the selection equations (see the references in Section 3.2). In the case of no endogeneity due to non-zero correlation between right-hand-side covariates and idiosyncratic errors, the approaches to tackle fixed effects in static panel data models with dichotomous selection are those proposed by Wooldridge (1995), Kyriazidou (1997) and Rochina-Barrachina (1999).[5] However, to our knowledge, no model of *polychotomous* choice that also allows for fixed effects in the selection and outcome equations has been considered in the literature. For panel models of polychotomous choice, but with no selection, see Honoré and Kyriazidou (2000a) or Magnac (2000).

We contribute to the literature by extending Wooldridge's (1995) estimator to the case of polychotomous selection in the spirit of Lee (1983) and applying this framework to estimate the returns to scale for all U.S. retail credit unions in 1994-2011. The latter has been recently brought into the spotlight of the scholarly discourse (Emmons & Schmid, 1999; Wilcox, 2005, 2006; Wheelock and Wilson, 2011). We compare our estimates to those (potentially biased and inconsistent) obtained by ignoring heterogeneity due to endogenous technology selection and unobserved effects.

Our main finding is that not all U.S. retail credit unions enjoy increasing returns to scale. When controlling for heterogeneity in the output mix, endogeneity in technology selection and unobserved effects among credit unions, we find that 6 to 12% of large credit unions offering all types of loans except commercial loans operate at decreasing returns to scale. These institutions should either reduce the scale of their operations or reconsider their output mixes. This finding, for instance, contrasts with the results in Wheelock and Wilson (2011) who find no significant evidence of decreasing returns to scale among all credit unions in their sample. We consistently fail to reject the null of exogenous technology selection among credit unions and generally find that ignoring endogeneity of this process or ignoring unobserved time-invariant effects across units leads to downward biases in returns to scale estimates.

However, having addressed all the concerns we raise in this paper, we find that the majority of credit unions (among all technology types) shows the evidence of substantial economies of

[5] Other methods consider dynamic panels with binary selection or allow for endogenous covariates (e.g., Vella & Verbeek, 1999; Semykina & Wooldridge, 2011; Kyriazidou, 2001; Semykina & Wooldridge, 2010; Charlier, Melenberg & von Soest, 2001; Dustmann & Rochina-Barrachina, 2007, where the last four are the extensions of the three papers cited in the main text).

scale which leads us to conclude that growth of the industry is far from reaching its peak. The industry-wide trends such as the diversification of financial services offered to members as well as mergers among credit unions are likely to persist over coming years. We therefore expect a policy debate over credit unions' tax-exempt status as well as their special regulatory treatment compared with commercial banks to reignite in the near future. As these institutions grow in size and complexity, they may become of systemic importance for regulators and the economy.

The rest of the paper proceeds as follows. Section 2 provides a description of the data as well as a discussion of how we identify heterogeneous credit union technologies. We describe our econometric model of polychotomous choice with endogenous selection and fixed effects in Section 3. Section 4 presents the results, and Section 5 concludes.

2 HETEROGENEOUS TECHNOLOGIES

In this section, we define the framework in which we examine credit union technologies. Due to their cooperative nature, credit unions are not profit-maximizers. Instead, they are thought of as maximizing service provision to their members in terms of quantity, price and variety of services (Smith, 1984; Fried *et al.*, 1999). Following a wide practice in the literature (Frame & Coelli, 2001; Frame *et al.*, 2003; Wheelock and Wilson, 2011, 2013), we adopt a "service provision approach" under which, given their production technologies,[6] credit unions minimize non-interest, variable cost subject to the levels and types of outputs, the competitive prices of variable inputs and the levels of quasi-fixed netputs.

We consider the following four outputs: real estate loans ($y1$), business and agricultural loans ($y2$), consumer loans ($y3$) and investments ($y4$). We further follow Frame *et al.* (2003) and Wheelock and Wilson (2011, 2013) and include two quasi-fixed netputs to capture the price dimension of the service provision by credit unions: the average interest rate on saving deposits ($\tilde{y}5$) and the average interest rate on loans ($\tilde{y}6$). The input dimension of credit union cost is captured by including the price of capital ($w1$) and the price of labor ($w2$). To partially account for the riskiness of the credit union operations, we also include equity capital ($\tilde{k}$) as a quasi-fixed input to the cost function, as usually done in the banking literature. The latter has been broadly taken for granted by the existing credit union literature under the implicit assumption of risk-neutral behavior of credit union managers. Including equity capital is also appropriate if one considers it as an additional input to the production of loans (see Hughes & Mester, 1998, 2011; Hughes, Lang, Mester & Moon, 1996; among many others). All of these variables are taken as arguments of the dual variable, non-interest cost function of a credit union, defined as

$$C(\boldsymbol{y}, \tilde{\boldsymbol{y}}, \boldsymbol{w}, \tilde{k}) = \min_{\boldsymbol{x}}\{\, \boldsymbol{x}'\boldsymbol{w} \mid T(\boldsymbol{y}, \boldsymbol{x}, \tilde{\boldsymbol{y}}, \tilde{k}) \leq 1;\ \tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}}_0;\ \tilde{k} = \tilde{k}_0 \,\}, \tag{2.1}$$

where $\boldsymbol{y} = (y1, y2, y3, y4)$ is a vector of outputs, $\tilde{\boldsymbol{y}} = (\tilde{y}5, \tilde{y}6)$ is a vector of quasi-fixed netputs with the corresponding vector of observed (fixed) values $\tilde{\boldsymbol{y}}_0$; and $\boldsymbol{w} = (w1, w2)$ is a vector of the

[6] That is, given the mix of financial services (outputs) that credit unions opt to provide to their members.

variable input prices; $\boldsymbol{x} = (x1, x2)$ is a vector of variable inputs; $\tilde{k}$ is a quasi-fixed input with the observed (fixed) value $\tilde{k}_0$; and $T(\boldsymbol{y}, \boldsymbol{x}, \tilde{\boldsymbol{y}}, \tilde{k})$ is the transformation function.

The data we use in this study come from year-end call reports available from the National Credit Union Administration (NCUA), a federal regulatory body that supervises credit unions. The available data cover all state and federally chartered U.S. credit unions over the period of 1994 to 2011. We discard observations with negative values of outputs and total cost. In addition, we exclude observations with non-positive values of variable input prices, quasi-fixed netputs, equity capital, total assets, reserves and total liabilities. Since $\tilde{\boldsymbol{y}}$ and $w1$ are interest rates, we also eliminate those observations for which values of these variables lie outside the unit interval. These excluded observations are likely to be the result of erroneous data reporting. For the details on construction of the variables from the call reports, see the Appendix.

In this paper we focus on retail, or so-called natural-person, credit unions only. We therefore exclude corporate credit unions (whose customers are the retail credit unions) from the sample in an attempt to minimize noise in the data due to apparent non-homogeneity between these two types of depositories (this results in a loss of 0.7% observations in the sample). Our data sample thus consists of 151,817 year-observations for all retail state and federally chartered credit unions over 1994-2011.

We next proceed to the identification of heterogeneous technologies among credit unions in our sample. As we pointed out in the Introduction, the data indicate the presence of significant differences among credit unions in terms of the mix of services they offer to members. Based on the tabulation of zero-value observations reported in Table 1, on average, we find that 88% of credit unions in our sample do not offer business loans ($y2$) and 31% do not offer mortgage loans ($y1$) in a given year. Ignoring this observed heterogeneity in the provision of services across credit unions amounts to making a strong assumption that all credit unions share the same technology that is invariant to the range of provided services. If the choice of the output mix is endogenous to credit unions' decisions, this assumption is unlikely to hold.

[Insert Table 2 here]

Given the four types of loans we consider in this paper, there are 15 possible technologies associated with unique output mixes which can be identified among credit unions. The possible heterogeneous technologies are those of the credit unions specialized in one (complete specialization), two or three types of loans (partial specialization) and of the unions that produce all four outputs (no specialization). Table 2 presents a summary of these technologies corresponding to output mixes constructed based on the non-zero-value loans reported by credit unions. The table shows that the majority of credit unions falls into the following three categories: (i) those that provide consumer loans and investments ($y3$, $y4$); (ii) those that provide real estate and consumer loans as well as investments ($y1$, $y3$, $y4$); and (iii) those that provide all types of outputs: real estate, business and consumer loans, and investments ($y1$, $y2$, $y3$, $y4$). Together, the three groups of credit unions constitute 99% of all observations in the sample, suggesting that the remaining one percent likely contains either outliers or reporting errors. We omit them from our analysis from this point forward. We label the three above output mixes as "1", "2" and "3", respectively, and define their corresponding technologies as "Technology 1", "Technology 2" and "Technology 3". We hereafter use technology and output mix types interchangeably when referring to credit unions.

[Insert Figure 1 here]

Figure 1 shows the breakdown of credit unions in our sample by their technology type. This figure indicates several trends. First, there is an apparent secular decline in the number of credit unions over time, mainly due to mergers and acquisitions. Second, the heterogeneity among U.S. credit unions (as captured by the technology type) is highly persistent over time. While today most credit unions still operate under Technology 2 as they did back in 1994, the presence of other technology types has increased over these years. Third, there is a trend among credit unions to shift away from Technology 1 to Technology 2 and even more so to Technology 3 (as confirmed by an unreported analysis of technology transitions).

[Insert Table 3 here]

To confirm that the credit unions belonging to heterogeneous technology types are intrinsically different from one another, consider Table 3 which presents summary statistics of the variables used in the dual cost function as well as several other variables descriptive of the characteristics of credit unions such as total assets, reserves, etc. (we will discuss the use of them in Section 4). All nominal stock variables are deflated to 2011 U.S. dollars using the GDP Implicit Price Deflator. A comparison of sample mean and median estimates of variables shows clear differences among technologies. As expected, the size of the credit unions (proxied either by total assets, reserves or the number of members) increases as one moves from Technology 1 to Technology 3. This is also apparent in Figure 2 which plots kernel density estimates for the log of total assets tabulated by technology types. The large differences between technology types favor our view that the assumption of homogeneous technology across credit unions is likely to result in the loss of information and the misspecification of the econometric model. As we show in Section 4, this produces biased estimates and potentially misleading results.

[Insert Figure 2 here]

3 A GENERALIZED MODEL OF CREDIT UNION TECHNOLOGIES

This section develops an econometric model that we employ in order to investigate underlying differences in heterogeneous technologies of U.S. credit unions. The model (i) avoids imposing a strong assumption of homogenous technology uniformly adopted by all credit unions irrespective of the service mix they offer to their members; (ii) explicitly accounts for the endogeneity of the selection of these different technologies by unions over the course of time; and (iii) allows for unobserved time-invariant fixed effects amongst credit unions.

3.1. A BASELINE MODEL

Consider a dual cost function of a credit union *i*:

$$C_i^s = g_s(\boldsymbol{x}_i^s; \boldsymbol{\beta}^s) + u_i^s \qquad \forall\ s = 1, \dots, S;\ i = 1, \dots, N \tag{3.1a}$$

$$T^{*s}_i = \boldsymbol{z}_i \boldsymbol{\gamma}^s + e_i^s\ , \tag{3.1b}$$

where C_i^s is the total variable, non-interest cost; $g_s(\cdot)$ is a *linear* (in parameters) cost function;[7] $\boldsymbol{x}_i^s$ is a $K_s \times 1$ vector of relevant variables as defined in Section 2; $\boldsymbol{\beta}^s$ and $\boldsymbol{\gamma}^s$are conformable parameter vectors. The superscript *s* denotes the technology type.

Note that C_i^s is observed only if the *s*th technology is chosen; T^{*s}_i is a latent variable determining the technology selection, given an $L \times 1$ vector of some relevant variables $\boldsymbol{z}_i$ which includes a vector of ones for the intercept (we define them in Section 4). While we later assume that the error terms u_i^s and e_i^s are orthogonal to $\boldsymbol{x}_i^s$ (under the premise of cost minimization) and $\boldsymbol{z}_i$, their distributions are however allowed to be correlated, namely $E[u_i^s e_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i] \neq 0$. Note that the above model is an extension of a standard endogenous selection model, also referred to as a switching regression model (Maddala, 1983, p. 223), to a case of polychotomous choice (Lee, 1978, 1982, 1983, 1995; Trost & Lee, 1984).

It is natural to think of the latent variable T^{*s}_i as measuring a credit union's propensity to select the technology *s*. The technology *s* is selected if and only if

$$T^{*s}_i > T^{*j}_i \qquad \forall\ j = 1, \dots, S\ \ (j \neq s) \tag{3.2}$$

While one can treat the switching among technologies as a system of $(S-1)$ dichotomous decision rules (Hay, 1980), we follow an alternative approach by considering the technology selection problem in McFadden's (1974) random utility framework, as suggested by Maddala (1983, p. 275) and Lee (1983). That is, the *s*th technology is said to be selected if and only if

$$T^{*s}_i > \max_{\substack{j=1,\dots,S \\ j \neq s}} \left\{T^{*j}_i\right\} \tag{3.3}$$

Define a categorical variable T such that $T_i = s$ if the *i*th credit union selects the technology *s*. Then, (3.3) can be rewritten as

$$T_i = s \quad \Leftrightarrow \quad \boldsymbol{z}_i \boldsymbol{\gamma}^s + e_i^s > \max_{\substack{j=1,\dots,S \\ j \neq s}} \left\{\boldsymbol{z}_i \boldsymbol{\gamma}^j + e_i^j\right\} \tag{3.4}$$

For convenience, let

$$\varepsilon_i^s = \max_{\substack{j=1,\dots,S \\ j \neq s}} \left\{\boldsymbol{z}_i \boldsymbol{\gamma}^j + e_i^j\right\} - e_i^s \tag{3.5}$$

Then it follows from (3.4) that

$$T_i = s \quad \Leftrightarrow \quad \varepsilon_i^s < \boldsymbol{z}_i \boldsymbol{\gamma}^s \tag{3.6}$$

[7] In this paper, we consider the translog cost function. For more on the choice of this specification, see Section 4.

In order to proceed further, we need to make a distributional assumption about the error term in the selection equation. The assumption we make is standard to polychotomous choice models.

> ASSUMPTION 1. For $s = 1, \dots, S$ and $i = 1, \dots, N$, the error term e_i^s is independent of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$, which implies $E[e_i^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i] = 0$, and is identically and independently distributed with the type I extreme-value distribution.

Then, it can be shown that (Domencich & McFadden, 1975)

$$\Pr[T_i = s] = \Pr[\varepsilon_i^s < \boldsymbol{z}_i\boldsymbol{\gamma}^s] = F_s(\boldsymbol{z}_i\boldsymbol{\gamma}^s) = \frac{\exp(\boldsymbol{z}_i\boldsymbol{\gamma}^s)}{\sum_{j=1}^{S}\exp(\boldsymbol{z}_i\boldsymbol{\gamma}^j)} \ , \tag{3.7}$$

which yields a multinomial logistic ε_i^s with the corresponding marginal distribution $F_s(\cdot)$.

Also note that, for some strictly positive monotonic transformation $J_s(\cdot)$, condition (3.6) is equivalent to

$$T_i = s \quad \Leftrightarrow \quad J_s(\varepsilon_i^s) < J_s(\boldsymbol{z}_i\boldsymbol{\gamma}^s) \tag{3.8}$$

We can now look at the benchmark model in (3.1) as a *binary* choice (sample selection) model, for each given technology s (Maddala, 1983, p.276). That is, we can essentially replace the technology-selection equation (3.1b) for each $s = 1, \dots, S$ with its equivalent

$$\tilde{T}^{*s}_i = J_s(\boldsymbol{z}_i\boldsymbol{\gamma}^s) - J_s(\varepsilon_i^s) \ , \tag{3.9}$$

where $\tilde{T}^{*s}_i$ is a transformed latent variable such that $T_i = s$ if and only if $\tilde{T}^{*s}_i > 0$, i.e., condition (3.8) is satisfied. Following Lee (1982, 1983), we consider $J_s(\cdot) \equiv \Phi^{-1}[F_s(\cdot)]$, where $\Phi(\cdot)$ is the standard normal cdf. The advantage of such transformation is that the random error $J_s(\varepsilon_i^s)$ in (3.9) is standard normal by construction, which would later enable us to employ well-known results on the truncated moments of the standard normal. For convenience, we define $\tilde{\varepsilon}_i^s \equiv J_s(\varepsilon_i^s)$.

Thus, a natural way to proceed is to specify a form of correlation between two disturbances in (3.1a) and (3.9), for each $s = 1, \dots, S$, which would permit the correction for selection bias in the outcome equation.

> ASSUMPTION 2. For $s = 1, \dots, S$ and $i = 1, \dots, N$, the error u_i^s is orthogonal to $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ and its conditional mean, given $\tilde{\varepsilon}_i^s$, is linear
>
> $$E[u_i^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i, \tilde{\varepsilon}_i^s] = E[u_i^s|\tilde{\varepsilon}_i^s] = L[u_i^s|\tilde{\varepsilon}_i^s] \ , \tag{3.10}$$
>
> where $L[\cdot]$ denotes the linear projection operator.

Specifically, we set

$$L[u_i^s|\tilde{\varepsilon}_i^s] = \pi^s\tilde{\varepsilon}_i^s \tag{3.11}$$

Olsen (1980) proposes the same assumption of a linear conditional mean of u_i^s in the binary selection setting, in order to derive a selection bias correction term. Maddala (1983, p. 269) parameterizes $\pi^s = \frac{\rho^s \sigma_u^s}{\sigma_{\tilde{\varepsilon}}^s}$ and then normalizes $\sigma_{\tilde{\varepsilon}}^s$ to unity thus setting $\pi^s = E[u_i^s \tilde{\varepsilon}_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i]$.[8] Also, a common alternative to Assumption 2 is the assumption of bivariate normality of the two disturbances which also implies linearity of the conditional mean of u_i^s [as in Lee (1982, 1983)]; however, our assumption is less restrictive.

Consider now the conditional mean of C_i^s, given that the sth technology is selected. From (3.1a) and (3.9) we get

$$\begin{aligned} E[C_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i, T_i = s] &= g_s(\boldsymbol{x}_i^s; \boldsymbol{\beta}^s) + E[u_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i, T_i = s] \\ &= g_s(\boldsymbol{x}_i^s; \boldsymbol{\beta}^s) + \pi^s E[\tilde{\varepsilon}_i^s | \tilde{\varepsilon}_i^s < J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)] \ , \end{aligned} \tag{3.12}$$

where we have made use of (3.8), (3.11) and the assumption of exogeneity of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ in the second equality. The truncated mean $E[\tilde{\varepsilon}_i^s | \tilde{\varepsilon}_i^s < J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)]$ has a known form of the negative of the inverse Mills ratio. Thus, from (3.12) we get

$$E[C_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i, T_i = s] = g_s(\boldsymbol{x}_i^s; \boldsymbol{\beta}^s) - \pi^s \lambda_s [J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)] \tag{3.13}$$

Here $\lambda_s[\cdot]$ denotes the inverse Mills ratio, i.e., $\lambda_s \equiv \frac{\phi[J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)]}{\Phi[J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)]} = \frac{\phi[J_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)]}{F_s(\boldsymbol{z}_i \boldsymbol{\gamma}^s)}$, where $\phi(\cdot)$ is the standard normal pdf.

The model can now be consistently estimated in two stages. The first-stage estimates of $\boldsymbol{\gamma}^s$ can be obtained via multinomial logit performed on (3.7), which are then used to compute the selection bias correction term $\hat{\lambda}_s[\cdot]$. Under the assumption of a linear (in parameters) functional form for $g_s(\boldsymbol{x}_i^s; \boldsymbol{\beta}^s)$, the consistent estimates of the main parameters of interest, namely $\boldsymbol{\beta}^s$, are obtained in the second stage via performing OLS on (3.13) that includes predicted $\hat{\lambda}_s[\cdot]$ in place of $\lambda_s[\cdot]$ (for each technology s, separately).

3.2. A Generalized Model with Fixed Effects

We now consider a generalized model of heterogeneous technologies under endogenous selection which also allows for unobserved credit-union-specific heterogeneity that is correlated with covariates in both the selection and the outcome equations. Thus, the benchmark model in (3.1) can be modified as

$$C_{it}^s = g_s(\boldsymbol{x}_{it}^s; \boldsymbol{\beta}^s) + \alpha_i^s + u_{it}^s \qquad \forall\ s = 1, \dots, S;\ i = 1, \dots, N;\ t = 1, \dots, t_{max} \tag{3.14a}$$

$$T^{*s}_{it} = \boldsymbol{z}_{it} \boldsymbol{\gamma}^s + \xi_i^s + e_{it}^s \ , \tag{3.14b}$$

where (α_i^s, ξ_i^s) are time-invariant, credit-union-specific fixed effects which are also specific to the choice of technology. We also add the time subscript t to differentiate between the cross-sectional and temporal variations. Note a notational change from now onward: (i) $\boldsymbol{x}_{it}^s$ and $\boldsymbol{z}_{it}$ denote $K_s \times 1$ and $L \times 1$ vectors of covariates at time t, respectively; (ii) $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ are now redefined as $\boldsymbol{x}_i^s \equiv$

---

[8] Note that neither Olsen (1980) nor Maddala (1983) transform the error in the selection equation (due to a binary nature of switching), instead they directly assume its normality.

$(\boldsymbol{x}_{i1}^s, \dots, \boldsymbol{x}_{it}^s, \dots, \boldsymbol{x}_{it_{max}}^s)$ and $\boldsymbol{z}_i \equiv (\boldsymbol{z}_{i1}, \dots, \boldsymbol{z}_{it}, \dots, \boldsymbol{z}_{it_{max}})$, respectively. All remaining variables and parameters are defined as before.

The estimation of a generalized model described by system (3.14) is not trivial. While there has been a great interest in extending traditional limited dependent variable models to the case of panel data which permits controlling for unobserved effects,[9] the literature on such models incorporated into linear regressions with selectivity mainly focuses on binary selection (for a comprehensive review, see Baltagi, 2008). These panel data selection models differ in their assumptions about the form of the unobserved heterogeneity in outcome and selection equations: whether random effects are assumed in both equations (Hausman & Wise, 1979; Ridder, 1990, 1992; and Verbeek & Nijman, 1996) or a combination of random and fixed effects is modeled (Verbeek, 1990). Few attempts have been made to allow for fixed-effect type heterogeneity in both outcome and selection equations. In the case of no endogeneity due to non-zero correlation between right-hand-side covariates and idiosyncratic errors, the three approaches to tackle fixed effects in these types of econometric models are those of Wooldridge (1995), Kyriazidou (1997) and Rochina-Barrachina (1999). These three papers mainly consider Type 2 Tobit model [Amemiya's (1985) terminology], whereas Wooldridge (1995) also explicitly talks of Type 3 Tobit. The extension of Kyriazidou's (1997) estimator to all types of Tobit (1 to 5) is discussed in Honoré and Kyriazidou (2000b). For a concise comparison of the three estimators, see Dustmann and Rochina-Barrachina (2007).

Given the research question that we posit in this paper, the model that we consider is of polychotomous choice with fixed effects in selection and outcome equations. To our knowledge, no such model has been considered in the literature. We thus fill in this void by extending Wooldridge's (1995) estimator to the case of polychotomous selection.

We first formalize the correlation between the fixed effects and the covariates in the selection equation (3.14b).

ASSUMPTION 1′. For $s = 1, \dots, S$ and $i = 1, \dots, N$, the fixed effects ξ_i^s in a selection equation s depend on the time averages of right-hand-side covariates, i.e.

$$\xi_i^s = \bar{\boldsymbol{z}}_i \boldsymbol{\eta}^s + c_i^s \tag{3.15a}$$

$$E[c_i^s | \boldsymbol{x}_i^s, \boldsymbol{z}_i] = 0 \ , \tag{3.15b}$$

where $\bar{\boldsymbol{z}}_i$ is a $(L-1) \times 1$ vector of time averages of $\boldsymbol{z}_{it}$ (excluding the unity vector); and $\boldsymbol{\eta}^s$ is a conformable vector of parameters. Substituting (3.15a) into (3.14b) yields

$$T^{*s}_{it} = \bar{\boldsymbol{z}}_i \boldsymbol{\eta}^s + \boldsymbol{z}_{it} \boldsymbol{\gamma}^s + v_{it}^s \ , \tag{3.15c}$$

[9] See, for instance, Chamberlain (1980), Manski (1987), Avery, Hansen and Hotz (1983), Sickles and Taubman (1986), Honoré and Kyriazidou (2000a), Honoré and Lewbel (2002), Magnac (2004).

where $v_{it}^s = c_i^s + e_{it}^s$ is independent of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ (which implies $E[v_{it}^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i] = 0$) and is identically and independently distributed with the type I extreme-value distribution over $i = 1, \ldots, N$ and $t = 1, \ldots, t_{max}$.[10]

Assumption 1′ warrants two remarks. First, the parameterization of fixed effects as in (3.15a) is quite popular in panel data models with limited dependent variables (see Mundlak, 1978; Nijman & Verbeek, 1992; Zabel, 1992).[11] It is a special case of a more general parameterization considered in Wooldridge (1995) who proposes estimation of the reduced form of the selection equation with time-varying parameters,[12] which can be expressed (in our notation) as

$$T^{*s}_{it} = \boldsymbol{z}_{i1}\boldsymbol{\delta}_{t1}^s + \cdots + \boldsymbol{z}_{it_{max}}\boldsymbol{\delta}_{tt_{max}}^s + v_{it}^s \tag{3.16}$$

While we acknowledge that (3.15c) is more restrictive than (3.16), we still opt for it due to its parsimony and computational simplicity. In particular, (3.16) would require estimation of $S[(L-1)t_{max} + 1]$ parameters for each time period *t*. In this study, given high nonlinearity of the objective function (the multinomial logit log-likelihood), the true values of $\boldsymbol{\delta}_{tt}^s$ may not be easy to locate. Further, the dataset we use is an unbalanced panel, which brings yet another set of complications since the number of parameters in (3.16) changes with each credit union.

Second, unlike Wooldridge (1995) who assumes normally distributed errors v_{it}^s, the distributional assumption we make is Gumbel, which is dictated by a polychotomous nature of the choice set. Also, v_{it}^s is assumed to be i.i.d over *i* and *t*, while Wooldridge (1995) specifies no temporal dependence in errors of the selection equation.[13] For instance, a similar "i.i.d over *i* and *t*" assumption in the case of the panel multinomial logit is made in Honoré and Kyriazidou (2000a) and Wooldridge (2010, p.653).

Following our steps from the previous subsection, the *s*th technology is said to be selected by a credit union *i* in the time period *t* if and only if

$$T^{*s}_{it} > \max_{\substack{j=1,\ldots,S \\ j\neq s}} \left\{T^{*j}_{it}\right\} \tag{3.17}$$

Redefine variable T so that $T_{it} = s$ if a credit union *i* switches to technology *s* in the time period *t*, and let [analogous to (3.5)]

$$\varepsilon_{it}^s = \max_{\substack{j=1,\ldots,S \\ j\neq s}} \left\{\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^j + \boldsymbol{z}_{it}\boldsymbol{\gamma}^j + v_{it}^j\right\} - v_{it}^s \ , \tag{3.18}$$

where we have made use of (3.15c). With this, we obtain [analogous to (3.6)]

$$T_{it} = s \quad \Leftrightarrow \quad \varepsilon_{it}^s < \bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s \tag{3.19}$$

---

[10] Since both $\bar{\boldsymbol{z}}_i$ and $\boldsymbol{z}_{it}$ contain unities, clearly it is only the sum of intercepts in $\boldsymbol{\eta}^s$ and $\boldsymbol{\gamma}^s$ that is identified.

[11] Note that such approach to a fixed-effect type of heterogeneity is also widely referred to as "correlated random effects", which purely is a matter of terminology. In this paper, we follow Wooldridge (1995), whose method we extend, in calling the effects "fixed".

[12] See his Assumption 2 (p.124).

[13] We partly compensate for this by including the time trend in the set of covariates $\boldsymbol{z}_{it}$.

Given that v_{it}^{s} is i.i.d with a Gumbel distribution, it follows that ε_{it}^{s} is multinomial logistically distributed over i and t with the corresponding marginal distribution $F_s(\cdot)$

$$\Pr[T_{it} = s] = F_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s) = \frac{\exp(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s)}{\sum_{j=1}^{S}\exp(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s)} \tag{3.20}$$

Similar to the cross-sectional case before, we treat a polychotomous switching model as a set of dichotomous choice models for each technology s. Using the $J_s(\cdot)$ transformation, we can replace (3.14b) for each $s = 1, \dots, S$ with its equivalent [analogous to (3.9)]

$$\tilde{T}^{*s}_{it} = J_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s) - J_s(\varepsilon_{it}^{s}) \ , \tag{3.21}$$

where $\tilde{T}^{*s}_{it}$ is a transformed latent variable such that $T_{it} = s$ if and only if $\tilde{T}^{*s}_{it} > 0$, i.e. the following equivalent of condition (3.19) is satisfied

$$T_{it} = s \quad \Leftrightarrow \quad J_s(\varepsilon_{it}^{s}) < J_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s) \tag{3.22}$$

We next make an assumption about the form of correlation between two disturbances in (3.14a) and (3.21) which enables us to correct for selection bias in outcome equations.[14] For convenience, we define $\tilde{\varepsilon}_{it}^{s} \equiv J_s(\varepsilon_{it}^{s})$.

ASSUMPTION 2′. For $s = 1, \dots, S$ and $i = 1, \dots, N$, the following holds for the sth outcome equation

(i) $E[u_{it}^{s}|\boldsymbol{x}_i^s, \boldsymbol{z}_i, \tilde{\varepsilon}_{it}^{s}] = E[u_{it}^{s}|\tilde{\varepsilon}_{it}^{s}] = L[u_{it}^{s}|\tilde{\varepsilon}_{it}^{s}]$ (3.23a)

(ii) $E[\alpha_i^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i, \tilde{\varepsilon}_{it}^{s}] = L[\alpha_i^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i, \tilde{\varepsilon}_{it}^{s}]$ (3.23b)

The first equality in Assumption 2′ (i) states that the error u_{it}^{s} is mean independent of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ conditional on $\tilde{\varepsilon}_{it}^{s}$. This assumption holds if u_{it}^{s} and $\tilde{\varepsilon}_{it}^{s}$ are independent of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ — a standard assumption made in the sample selection models (which we have also made in the previous section, partly motivating by the cost minimization behavior). Unlike Wooldridge (1995), we condition the expectation of u_{it}^{s} on $\boldsymbol{z}_i$ as well. This is necessary because we allow outcome and selection equations to have different covariates and non-zero (cross-equation) correlation between fixed effects. Further, note that (3.23a) does not impose any restrictions on temporal dependence of u_{it}^{s} or on the relationship between u_{it}^{s} and $\tilde{\varepsilon}_{it}^{s}$.

Similar to (3.11), we specify the linear projection of u_{it}^{s} on $\tilde{\varepsilon}_{it}^{s}$ in (3.23a) as

$$L[u_{it}^{s}|\tilde{\varepsilon}_{it}^{s}] = \pi_t^s\tilde{\varepsilon}_{it}^{s} \ , \tag{3.24}$$

where the parameter π_t^s is now allowed to be time-varying, thus emphasizing the presence of temporal dynamics in the relationship between u_{it}^{s} and $\tilde{\varepsilon}_{it}^{s}$.

In order to account for fixed effects in outcome equations, Assumption 2′ (ii) specifies the structure of unobserved heterogeneity. One can consider a general form of (3.23b) such as

$$L[\alpha_i^s|\boldsymbol{x}_i^s, \boldsymbol{z}_i, \tilde{\varepsilon}_{it}^{s}] = \boldsymbol{x}_{i1}^s\boldsymbol{\varphi}_{t1}^s + \dots + \boldsymbol{x}_{it_{max}}^s\boldsymbol{\varphi}_{tt_{max}}^s + \boldsymbol{z}_{i1}\boldsymbol{\omega}_{t1}^s + \dots + \boldsymbol{z}_{it_{max}}\boldsymbol{\omega}_{tt_{max}}^s + \psi_t^s\tilde{\varepsilon}_{it}^{s} \tag{3.25}$$

---

[14] For a counterpart in Wooldridge (1995), see his Assumption 3′ on p.126.

However, using the law of iterated expectations, one can easily show that under Assumptions (1′) and (2′) the parameters on $\boldsymbol{x}_{it}^s$ and $\boldsymbol{z}_{it}$ in (3.25) are necessarily constant over t.[15] Thus, (3.23b) simplifies to[16]

$$E[\alpha_i^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,\tilde{\varepsilon}_{it}^s] = \boldsymbol{x}_{i1}^s\boldsymbol{\varphi}_1^s + \cdots + \boldsymbol{x}_{it_{max}}^s\boldsymbol{\varphi}_{t_{max}}^s + \boldsymbol{z}_{i1}\boldsymbol{\omega}_1^s + \cdots + \boldsymbol{z}_{it_{max}}\boldsymbol{\omega}_{t_{max}}^s + \psi_t^s\tilde{\varepsilon}_{it}^s$$
$$= \boldsymbol{x}_i^s\boldsymbol{\varphi}^s + \boldsymbol{z}_i\boldsymbol{\omega}^s + \psi_t^s\tilde{\varepsilon}_{it}^s \qquad (3.26)$$

We are now ready to proceed to the derivation of the selection bias corrected cost function that also controls for unobserved effects. Taking the conditional mean of C_{it}^s from (3.14a), we obtain

$$\begin{aligned} E[C_{it}^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,T_{it}=s] &= g_s(\boldsymbol{x}_{it}^s;\boldsymbol{\beta}^s) + E[\alpha_i^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,T_{it}=s] + E[u_{it}^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,T_{it}=s] \\ &= g_s(\boldsymbol{x}_{it}^s;\boldsymbol{\beta}^s) + \boldsymbol{x}_i^s\boldsymbol{\varphi}^s + \boldsymbol{z}_i\boldsymbol{\omega}^s + (\psi_t^s+\pi_t^s)E[\tilde{\varepsilon}_{it}^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,T_{it}=s] \\ &= g_s(\boldsymbol{x}_{it}^s;\boldsymbol{\beta}^s) + \boldsymbol{x}_i^s\boldsymbol{\varphi}^s + \boldsymbol{z}_i\boldsymbol{\omega}^s + \\ &\quad + (\psi_t^s+\pi_t^s)E[\tilde{\varepsilon}_{it}^s|\tilde{\varepsilon}_{it}^s < J_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s)] \ , \end{aligned} \qquad (3.27)$$

where we have used (3.24) and (3.26) in the second equality, and (3.22) and the exogeneity of $\boldsymbol{x}_i^s$ and $\boldsymbol{z}_i$ in the last equality. Given that $\tilde{\varepsilon}_{it}^s$ is standard normal by construction, the expected value term in (3.27) equals the negative of the inverse Mills ratio. Therefore, the selection bias corrected outcome equation with unobserved heterogeneity simplifies to its final form

$$E[C_{it}^s|\boldsymbol{x}_i^s,\boldsymbol{z}_i,T_{it}=s] = g_s(\boldsymbol{x}_{it}^s;\boldsymbol{\beta}^s) + \boldsymbol{x}_i^s\boldsymbol{\varphi}^s + \boldsymbol{z}_i\boldsymbol{\omega}^s - \rho_t^s\,\lambda_s[J_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s + \boldsymbol{z}_{it}\boldsymbol{\gamma}^s)] \ , \qquad (3.28)$$

where $\rho_t^s \equiv (\psi_t^s + \pi_t^s)$ and $\lambda_s[\cdot] \equiv \frac{\phi[J_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s+\boldsymbol{z}_{it}\boldsymbol{\gamma}^s)]}{F_s(\bar{\boldsymbol{z}}_i\boldsymbol{\eta}^s+\boldsymbol{z}_{it}\boldsymbol{\gamma}^s)}$.

The generalized model is consistently estimated via a two-stage procedure. The first stage is the (pooled) multinomial logit with fixed effects as specified in (3.20), the estimates of which are then used to compute $\hat{\lambda}_s[\cdot]$. The second stage consists of estimating (3.28), in which predicted $\hat{\lambda}_s[\cdot]$ is used in place of $\lambda_s[\cdot]$, via pooled OLS (for each technology s, separately) after an assumption of a linear (in parameters) form for $g_s(\boldsymbol{x}_{it}^s;\boldsymbol{\beta}^s)$ is made. Note that the outcome equations now also include $\boldsymbol{z}_i$ which is a consequence of allowing selection and outcome equations to be determined by different sets of covariates.

4 ESTIMATION AND RESULTS

We estimate both the benchmark and generalized models described in Section 3. For this, we need to specify the set of variables $\boldsymbol{z}$ which enter selection equations (3.1b) and (3.14b) that govern endogenous switching between technology types by credit unions. These variables must be relevant to a credit union's decision about the range of services it seeks to offer to its members.[17] As noted in Section 2, the data particularly suggest considering covariates that correlate with the

[15] See Wooldridge (1995) for details.
[16] Note again that since both of $\boldsymbol{x}_{it}^s$ and $\boldsymbol{z}_{it}$ contain unities, the $2t_{max}$ intercept parameters in $\boldsymbol{\varphi}_t^s$ and $\boldsymbol{\omega}_t^s$ are not identified individually, although their sum is identified.
[17] Recall that we define technology types based on the output mixes endogenously chosen by credit unions.

size of a credit union such as its total assets and other variables reflecting credit union's financial strength and potential for growth. After carefully examining the existing literature for potential candidates, we settle on the following set of variables: total assets, reserves, leverage ratio,[18] the number of current and potential members, indicator variables for federally accredited, state accredited and federally insured,[19] and multiple-bond credit unions. Table 3 provides their summary statistics.

We use the total value of assets and the number of current members of the credit union to capture the size of credit unions (Goddard *et al*., 2002); one can naturally expect a larger credit union to seek the diversification of its output mix and thus switch to a less specialized technology. We proxy the credit union's potential for growth using the reported level of reserves (Bauer, 2008, 2009) and the size of the field of membership, i.e., the number of potential members (Goddard *et al*., 2008). The intuition here is as follow. The larger a credit union's field of membership is, the more likely it is to consider offering a wider range of services to its members and thus changing its technology. Similarly, the selection equation includes the leverage ratio to control for the level of financial constraint a credit union may be subject to, which can directly influence its growth. We also condition the choice of technology on whether a credit union can draw its members from a pool of people with single or multiple associations. This is crucial since multiple-bond credit unions have a substantial advantage over single-bond ones due to their ability to grow in size and diversify credit risks more easily (Walter, 2006). For instance, a single-bond credit union that is authorized to draw its members from a pool of employees of a single plant only is susceptible to any economic shock that this plant it subject to. Dummies for federally and state accredited credit unions are used to control for possible intrinsic differences between the two types of entities. We also include the vector of ones for the intercept term and the time trend to capture temporal dynamics in technology switching.[20]

In order to analyze the consequences of the failure to accommodate heterogeneity in technologies resulting from endogenous selection as well as the presence of unobserved effects amongst credit unions, we estimate several auxiliary models in addition to those developed in Section 3. For the ease of discussion, all the models we estimate are defined below.

Models Ignoring Unobserved Effects:

Model 1. The baseline model of heterogeneous technologies with endogenous switching; given by (3.1) and estimated in two stages as described in Section 3.1.

Model 2. The model of heterogeneous technologies under the assumption of exogenous (ignorable) switching; estimated via pooled OLS using (3.1a) separately for each technology. We estimate this model to investigate the degree with which results change if one does not

[18] Defined as the ratio of total debt to total assets.
[19] While all federally accredited unions are insured, the same however cannot be said about all state accredited unions.
[20] All continuous variables are logged to allow for some degree of nonlinearity as well as to scale down the values of covariates.

recognize the endogeneity of technology selection by credit unions. Any differences between Models 1 and 2 are attributed to selection bias in the latter.

Model 3. The model of homogeneous technology. This model is most widely estimated in the existing literature by specifying two outputs instead of four in order to eliminate zero-value observations. The two outputs are the linearly aggregated loans ($y1 + y2 + y3$) and investments ($y4$). The model is estimated via pooled OLS using the whole sample ignoring a credit union's technology type.

Models Accounting for Unobserved Effects:

Model 4. The generalized model of heterogeneous technologies with endogenous switching and fixed effects; given by (3.14) and estimated in two stages as described in Section 3.2.

Model 5. The model of heterogeneous technologies under the assumption of exogenous (ignorable) switching with fixed effects; estimated via pooled OLS using (3.28) with selection bias correction terms omitted (separate regressions for each technology type). In order to facilitate direct comparability between the models, here we account for fixed effects in the same fashion as in Model 4, i.e., by parameterizing the correlation between unobserved effects and the right-hand-side covariates in the spirit of Assumption 2′ (ii).[21]

Model 6. The model of homogeneous technologies with two outputs and fixed effects; estimated via pooled OLS using observations for credit unions of all technology types. Similar to the argument above, here we also control for fixed effects by parameterizing the correlation between unobserved effects and the right-hand-side covariates.

For all models, we use the translog form[22] of the dual cost function $g_s(\cdot)$, onto which we impose the symmetry and linear homogeneity (in input prices) restrictions. In the first-stage estimations of Models 1 and 4 (i.e., multinomial logit), parameter vectors $\boldsymbol{\eta}^s$ and $\boldsymbol{\gamma}^s$ are normalized to zero for $s = 3$. To conserve space, we do not report the results from the first stage (they are available upon request) and thus directly proceed to the discussion of the main results.

The left pane of Table 4 reports the summary statistics of the point estimates of returns to scale based on Models 1 through 3, over the entire sample period of 1994-2011. When computing these statistics, we omit the first and the last percentiles of the distribution of the returns to scale estimates, in order to minimize the influence of outliers. However, the omitted estimates

[21] An alternative would be to estimate Model 5 via the within estimator that assumes no form of correlation between fixed effects and regressors in the cost function equation. Instead, the fixed effects are parameterized in spirit of (3.26) as $E[\alpha_i^s|\boldsymbol{x}_i^s] = L[\alpha_i^s|\boldsymbol{x}_i^s] = \boldsymbol{x}_i^s\boldsymbol{\varphi}^s$. Here, we do not condition on the error or covariates from the selection equation since Model 5 ignores endogenous selectivity.

[22] While we emphasize the heterogeneity in credit unions' production technologies due to their differing output mixes, we acknowledge that ideally one would also prefer to allow the technology to be heterogeneous among credit unions for a *given* output mix. In this paper, we assume such heterogeneity away, which is an undeniable limitation of our analysis. One could extend our model to allow the cost function to be credit-union specific by, say, employing semi- or nonparametric methods (although controlling for fixed effects in that case may require a different approach). Here, we opt for the parametric specification (translog) mainly for expository purposes as well as its tractability. We leave the extension of our model to an even more general setup for future research.

correspond to the same observations across all six models, in order to keep the results comparable. We still can cross-reference results of different models on the credit union level.

[Insert Table 4 here]

In Table 4, we break down the results by the technology type of credit unions. Note that although Models 1 and 2 estimate credit unions' cost functions for each technology separately, we also report the statistics for the whole distribution of credit unions obtained by pooling the results (over technology types) after the estimation. Similarly, we are able to break down the estimates of returns to scale from Model 3 by technology types after fitting a single homogeneous cost function for all credit unions. The credit-union-specific estimates of returns to scale are obtained using the formula that takes into account the quasi-fixity of equity capital (Caves, Christensen & Swanson, 1981)

$$RS = \frac{1 - \frac{\partial \log C}{\partial \log \tilde{k}}}{\sum_j \frac{\partial \log C}{\partial \log y_j}} \ , \qquad (4.1)$$

where $y_j \in \boldsymbol{y}$ are the outputs a credit union produces.

The empirical evidence suggests that, when compared to a benchmark Model 1, Models 2 and 3 tend to underestimate the returns to scale across all three technology groups (more so for Technologies 1 and 3). One can see it in Figure 3 which plots kernel densities of the estimated returns to scale from these models. Biases (due to ignored selectivity and/or heterogeneity among technologies) in returns to scale estimates from Models 2 and 3 tend to be downward for credit unions operating under Technologies 1 and 3, whereas we cannot unambiguously claim the sign of these biases in the case of Technology 2.

[Insert Figure 3 here]

We perform a formal test for selection bias on the coefficient of the inverse Mills ratio in Model 1 [equation (3.13)], i.e., a t test of $\mathbb{H}_0$: $\pi^s = 0$ for $s = 1,2,3$. The tests reject the null of no selection bias with the p-values of less than 10^{-4} for all three technology groups, confirming that the switching is not exogenous and hence not "ignorable". The latter validates the proposition that the estimates from Models 2 and 3 are subject to selection bias. Similarly, we test the proposition of non-homogenous technologies across credit unions with different output mixes. The multiple-restriction Wald test of $\mathbb{H}_0$: $\boldsymbol{\beta}^s = \boldsymbol{\beta}^j$ for $s = 1,2,3$ $(s \neq j)$ on the coefficients of (3.13) strongly confirms the presence of heterogeneity in credit union cost structures: the p-value is less than 10^{-75}. Note that in order to conduct this inference we need to estimate the variance-covariance matrix for Model 1, which is complicated due to its two-stage estimation procedure. To overcome this complication, we follow Newey's (1984) suggestion and rewrite the (two-stage) model in a multiple-equation GMM framework which permits derivation of an asymptotic variance-covariance matrix of the estimator.[23] We perform a similar exercise when computing the variance-covariance matrix for the generalized Model 4.

The qualitative differences between the models are more transparent when credit unions are grouped into three returns to scale categories: decreasing returns to scale (DRS), constant return to scale (CRS) and increasing returns to scale (IRS). We classify a credit union as exhibiting DRS/CRS/IRS if the point estimate of its returns to scale is found to be statistically less than/equal to/greater than unity at the 95% significance level.[24]

[Insert Figure 4 here]

Figure 4 depicts the 95% confidence intervals of the returns to scale estimates from Model 1, based on which the right pane of Table 4 is partly populated.[25] These confidence intervals, which correspond to each observation (credit-union-year) over the 1994-2011 period, are represented by vertical line segments that are sorted by the lower bound. Based on the results from Model 1 (also see Table 4), we find that virtually all credit unions with Technologies 1 and 3 operate under IRS. We however cannot say the same with respect to credit unions operating under Technology 2. Here we find that 10,626 out of 85,381 credit-union-years (12.4%) exhibit DRS and 2,626 (3.1%) exhibit CRS.

It might seem at first glance that the results do not differ that much across the three models qualitatively, at least in the case of Technology 1. For this technology group, Models 2 and 3 classify 98.1% and 99.8% of the subsample, respectively, as operating at IRS (43,444 and 44,183 out of 44,274 credit-union-years, respectively). However, (unreported) Spearman's rank correlation coefficients of the returns to scale estimates between the three models reveal that there is an astonishingly weak, if any at all, correspondence in rankings of credit unions between Model 1 and Models 2 and 3 (the correlation coefficients of 0.035 and 0.105, respectively).

Both Models 1 and 2 however produce relatively similar results for the credit unions operating under Technology 3. The one that stands out here is Model 3, according to which 22.9% of the subsample (4,081 out of 17,757 credit-union-years) exhibits DRS (see Table 4). Expectedly, there is weak rank correlation (equals to 0.262) between the rankings of credit unions of this model and the benchmark Model 1. The only instance when Model 1 predicts a larger number of credit unions operating under DRS than the remaining two models is for Technology 2 (12% of observations).

[Insert Table 5 here]

---

[23] The estimated variance is robust to heteroskedasticity in the cost functions.
[24] We use the delta method to construct standard errors for the returns to scale estimates.
[25] Similar figures for the other two models are available upon request.

However, the above results are still likely to be misleading because of endogeneity bias due to the ignored fixed effects among credit unions. We thus proceed to the models that explicitly control for unobserved effects.[26]

[Insert Figure 5 here]

Figure 5 plots the kernel densities of the returns to scale estimates from Models 4 through 6 (see Table 5 for the summary statistics of these estimates). The evidence again suggests that the models which ignore endogenous switching (Models 5 and 6) tend to underestimate the returns to scale at which credit unions operate across all three technology groups. The kernel densities of estimates from Model 4 are generally shifted leftward compared to those of estimates from Models 5 and 6. Thus, the biases in returns to scale estimates produced by Models 5 and 6 generally appear to be of negative sign (with some ambiguity in the case of Technology 2).

The Wald tests of $\mathbb{H}_0$: $\rho_1^s = \cdots = \rho_t^s = \cdots = \rho_{t_{max}}^s = 0$ for $s = 1,2,3$ performed on (3.28) again confirm the presence of selection bias in Models 5 and 6 (the p-values of less than 10^{-90} for all three technology groups). Similarly, we again reject the null of homogenous cost function across different technology groups. The p-value corresponding to the Wald test of $\mathbb{H}_0$: $\boldsymbol{\beta}^s = \boldsymbol{\beta}^j$ for $s = 1,2,3$ $(s \neq j)$ on the coefficients of (3.28) is less than 10^{-100}.

[Insert Figure 6 here]

Figure 6 shows the differences between Models 4 through 6 that account for credit union-specific fixed effects and those that ignore this unobserved heterogeneity (Models 1 through 3). The figure plots kernel densities of the returns to scale estimated by all six models. The evidence indicates the presence of a negative bias in the returns to scales estimates obtained from Models 1 through 3: the kernel densities from these models are to the left of those produced by the corresponding models that control for unobserved effects. The biases appear to be the largest in the case of Technology 3. The shift in the estimated returns to scale of credit unions can also be seen in Figure 7 which plots the 95% confidence intervals of the returns to scale estimates from the generalized Model 4. The estimated intervals have shifted upwards for Technologies 2 and 3, compared to those plotted in Figure 4. The above result emphasizes the importance of taking unobserved effects into account when quantifying credit union technologies.

[Insert Figure 7 here]

As expected, Models 4 to 6 predict a smaller number of credit unions with non-IRS across all technology groups than Models 1 to 3, respectively (compare right panes of Tables 4 and 5). Several issues, however, warrant a further discussion. Although all three Models 4, 5 and 6 strongly support the evidence in favor of IRS almost universally exhibited by credit unions operating under Technology 1, the rankings of these credit unions is strikingly different across

[26] Following equation (3.28), we parameterize fixed effects in cost functions as linear projections of (i) all continuous variables included in the first-stage selection equation and (ii) all unique variables in the cost functions, except for the time trend. Thus, we do not include squared and cross-product terms from the translog cost functions into the set of variables onto which fixed effects are assumed to project. Doing the latter would be redundant.

these models (similar to the case of Models 1 through 3). In particular, we find that the (unreported) correlation coefficients of rankings of credit unions in terms of their returns to scale between Models 4 and Models 5 and 6 are weak (if not close to zero): 0.060 and 0.341, respectively. We attribute these differences to selection and misspecification biases present in Models 5 and 6. In the case of Technology 3, however, we find that both Models 4 and 5 produce similar results: virtually every single credit union enjoys IRS. Model 6 is the one that stands out. It predicts that 3,212 out of 17,757 credit-union-years (18.1%) still exhibit non-IRS.

We find most agreement in the results across the models in the case of Technology 2: the (unreported) rank correlation coefficients are the largest in this case (around 0.9). However, based on the estimated generalized Model 4, we still find a relatively large number of credit unions that exhibit DRS and CRS (4,836 and 2,133 out of 85,381 credit-union-years, respectively). Figure 9 shows that these credit unions are the largest in the group. The latter, for instance, contrasts the results in Wheelock and Wilson (2011) who generally find no evidence of DRS and CRS among credit unions in their sample. These differences can be attributed to several reasons. First, the sample periods differ: we consider the period of 1994-2011, whereas Wheelock and Wilson (2011) examine the 1989-2006 period. Second, Wheelock and Wilson (2011) obtain their returns to scale estimates from an admittedly more flexible nonparametric cost function whereas our estimation approach is parametric. Third, they aggregate outputs in order to eliminate zero-value observations, and their cost function does not include equity capital as one of the inputs. Fourth, Wheelock and Wilson (2011) do not explore the possibility of endogeneity in a credit union's choice of the output mix. Lastly, while controlling for time effects, Wheelock and Wilson (2011) however do not address the issue of unobserved time-invariant heterogeneity among credit unions in the panel. All of these issues can potentially result in differences between Wheelock and Wilson's (2011) and our results.

[Insert Figure 8 here]

We find unexpected results when analyzing the relationship between returns to scale of a credit union and its size (proxied by total assets). Normally, one would expect to see an inverse relationship between the two. We do confirm it when looking at the entire sample. However, as Figure 8 shows, this result is not uniform across all technology groups. We find that the estimated returns to scale do largely fall as one moves from small to larger credit unions that operate under Technology 2. However, there is hardly any change in returns to scale among credit unions in the third technology group. Moreover, the returns to scale appear to increase with the size for credit unions operating under Technology 1. For instance, the estimates of returns to scale from Models 5 and 6 fall with the asset size *regardless* of the technology (not reported to conserve space). While these findings look puzzling at the first glance, there is an intuitive explanation to them.

Recall that the asset size of the credit unions increases as one moves from Technology 1 to 3 (see Table 3 and Figure 2). Therefore, given that the credit unions which operate under Technology 1 are already small in size altogether, an increase in available resources as a credit union grows enables it to adopt new information processing technologies that are initially quite expensive to install but, once installed, are substantial cost-savers. The example of such technologies would be internet banking, automated teller machines, use of electronic money as well as an access to members' credit history through the credit rating bureaus. Given that 25% of credit unions in the first technology group are as small as an entity with only 1 full-time equivalent employee, many of them are not financially capable of adopting the abovementioned technologies until they grow in size. The impact of these financial constraints however wears out as credit unions continue to grow which we indeed observe in the case of entities that move to the operation under Technology 2.

A potential explanation why the relationship between the size and returns to scale breaks down for depositories in the third technology group is greater diversification enjoyed by these larger credit unions. On average, credit unions in this group have a four times larger number of members than those belonging to the second technology group (an average of 12,700 vs. 3,600 members, respectively). The diversification comes not only through a larger membership pool, but also through a larger range of services provided to members as well as an opportunity to engage in more advanced financial operations (Wilcox, 2005). The latter is partly due to economies of diversification enjoyed by credit unions as they move from one technology to another (recall that technologies are nested). The data suggest the presence of non-negligible economies of scope, which is a matter of substantial interest on its own. We leave the discussion of it for a future paper. Lastly, larger credit unions can also protect their market positions by erecting entry barriers thus partly mitigating the decline in returns to scale as they grow.

5 CONCLUSION

A trillion dollar worth credit union industry takes up a significant portion of the U.S. financial services market, catering to almost a hundred million people in the country. Given the dramatic growth of the industry over the past few decades, there has been a substantial interest in formally modeling the technologies of credit unions. However, the econometric approaches widely used in the existing literature somewhat limit our understanding of the structure, dynamics and future evolution of the credit union industry.

Faced by the presence of an overwhelming number of observations for which the reported values of credit unions' outputs are zeros in the data, the existing studies of credit union technologies have mainly resorted to the linear aggregation of different types of outputs into broader categories. We believe this procedure leads to a loss of information in both econometric and economic senses. The presence of zero-value observations is not merely a data issue but a consequence of substantial time-persistent heterogeneity amongst credit unions' technologies as captured by differing output mixes. This heterogeneity is likely to be an outcome of an endogenous choice made by credit unions. Models that *a priori* impose homogeneity and/or overlook credit unions' endogenous technology selection are likely to produce biased, inconsistent and thus misleading estimates. The results are also likely to be biased due to unobserved effects which are widely ignored in the credit union literature.

In this paper, we address the above concerns by developing a unified framework that allows estimation of credit union technologies that is robust to (i) misspecification due to an *a priori* assumption of homogenous technology, (ii) selectivity bias due to ignoring the endogeneity in technology selection, and (iii) endogeneity (omitted variable) bias due to a failure to account for unobserved union-specific effects that are correlated with covariates in the estimated equations. To accommodate the above concerns, we develop a generalized model of endogenous switching with polychotomous choice and fixed effects by extending Wooldridge's (1995) estimator. We note that the developed model is not tailored to the analysis of credit unions only. The framework can be applied to any other panel data study (with fixed effects) where polychotomous selection applies. Some examples would be studies of electric or water utilities, which often include both specialized and integrated companies that operate under non-homogeneous technologies.

Our main finding is that not all U.S. retail credit unions seem to uniformly enjoy increasing returns to scale. When controlling for heterogeneity in the output mix, endogeneity in technology selection and unobserved effects among credit unions, we find that a non-negligible number of large credit unions (6 to 12% of those offering all types of loans but commercial) operates at decreasing returns to scale. The latter implies that these institutions should potentially reduce their size or reconsider their output mix. We consistently fail to reject the null of exogenous technology selection among credit unions and generally find that ignoring endogeneity of this process produces negative biases in the estimates of returns to scale. We also document downward biases in the return to scale estimates when the model fails to account for unobserved time-invariant effects.

After addressing all the concerns we raise in this paper, we find that the majority of credit unions (among all technology types) operate under substantial economies of scale which leads us to conclude that the growth of the industry is far from reaching its peak. Thus, the industry-wide trends like the diversification of the range of financial services offered to members as well as mergers and acquisitions among credit unions are likely to persist over the coming years. Our results can therefore contribute to the policy debate over credit unions' tax-exempt status as well as their special regulatory treatment compared with commercial banks. As these institutions grow in size and complexity, they may become of systemic importance for regulators and the economy.

REFERENCES

Amemiya, T. (1985). *Advanced Econometrics*. Cambridge: Harvard University Press.

Avery, R.B., Hansen, L.P., Hotz, V.J. (1983). Multiperiod Probit Models with Orthogonality Condition Estimation. *International Economic Review*, 24, 21-35.

Baltagi, B. (2008). *Econometric Analysis of Panel Data*. Wiley, John & Sons Inc.

Bauer, K. (2008). Detecting Abnormal Credit Union Performance. *Journal of Banking and Finance*, 32, 573-586.

Bauer, K.J., Miles, L.L., Nishikawa, T. (2009). The Effect of Mergers on Credit Union Performance. *Journal of Banking and Finance*, 33, 2267-2274.

Caves, D.W., Christensen, L.R., Swanson, J.A. (1981). Productivity Growth, Scale Economies, and Capacity Utilization in U.S. Railroads, 1955-1974. *American Economic Review*, 71 (5), 994-1002.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *Review of Economic Studies*, 47 (1), 225-238.

Charlier, E., Melenberg, B., von Soest, A. (2001). An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data. *Journal of Econometrics* 101, 71-107.

Domencich, T.A., McFadden, D. (1975). *Urban Travel Demand*. Amsterdam: North-Holland Publishing.

Dustmann, C., Rochina-Barrachina, M.E. (2007). Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations. *Econometrics Journal*, 10, 263-293.

Emmons, W.R., Schmid, F.A. (1999). Credit Unions and the Common Bond. *Federal Reserve Bank of St. Louis Review*, 81, 41-64.

Frame, W.S., Coelli, T.J. (2001). U.S. Financial Services Consolidation: The Case of Corporate Credit Unions. *Review of Industrial Organization*, 18, 229-242.

Frame, W.S., Karels, G.V., McClatchey, C.A. (2003). Do Credit Unions Use Their Tax Advantage to Benefit Members? Evidence from a Cost Function. *Review of Financial Economics*, 12, 35-47.

Fried, H.O., Lovell, C.A.K., Yaisawarng, S. (1999). The Impact of Mergers on Credit Union Service Provision. *Journal of Banking and Finance*, 23, 367-386.

Goddard, J.A., McKillop, D.G., Wilson, J.O.S. (2002). The Growth of US Credit Unions. *Journal of Banking and Finance*, 26, 2327-2356.

Goddard, J.A., McKillop, D.G., Wilson, J.O.S. (2008). The Diversification and Financial Performance of US Credit Unions. *Journal of Banking and Finance*, 32, 1836-1849.

Hausman, J.A., Wise, D. (1979). Attrition Bias in Experimental and Panel Data: the Gary Income Maintenance Experiment. *Econometrica*, 47, 455-473.

Hay, J. (1980). Selectivity Bias in a Simultaneous Logit-OLS Model. *Manuscript*. University of Connecticut Health Center.

Honoré, B.E., Kyriazidou, E. (2000a). Panel Data Discrete Choice Models with Lagged Dependent Variables. *Econometrica*, 68 (4), 839-874.

Honoré, B.E., Kyriazidou, E. (2000b). Estimation of Tobit-type Models with Individual Specific Effects. *Econometric Reviews*, 19 (3), 341-366.

Honoré, B.E., Lewbel, A. (2002). Semiparametric Binary Choice Panel Data Model without Strictly Exogenous Regressors. *Econometrica*, 70, 2053-2063.

Hughes, J.P, Mester, L.J. (1998). Bank Capitalization and Cost: Evidence of Scale Economies in Risk Management and Signaling. *Review of Economics and Statistics*, 80 (2), 314-325.

Hughes, J.P, Mester, L.J. (2011). Who Said Large Banks Don't Experience Economies of Scale? Evidence from a Risk-Return-Driven Cost Function. *Working Paper* No. 11-27, Federal Reserve Bank of Philadelphia.

Hughes, J.P., Lang, W., Mester, L.J., Moon, C. (1996). Efficient Banking Under Interstate Branching. *Journal of Money, Credit and Banking*, 28 (4), 1045-1071.

Kyriazidou, E. (1997). Estimation of a Panel Data Sample Selection Model. *Econometrica*, 65 (6), 1335-1364.

Kyriazidou, E. (2001). Estimation of Dynamic Panel Data Sample Selection Models. *Review of Economic Studies*, 68, 543-572.

Lee, L.F. (1978). Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables. *International Economic Review*, 19 (2), 415-433.

Lee, L.F. (1982). Some Approaches to the Correction of Selectivity Bias. *Review of Economic Studies*, 49 (3), 355-372.

Lee, L.F. (1983). Generalized Econometric Models with Selectivity. *Econometrica*, 51 (2), 507-512.

Lee, L.F. (1995). The Computation of Opportunity Costs in Polychotomous Choice Models with Selectivity, *Review of Economics and Statistics*, 77 (3), 423-435.

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

Magnac, T. (2000). Subsidised Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories. *Economic Journal*, 110, 805-837.

Magnac, T. (2004). Panel Binary Variables and Sufficiency: Generalizes Conditional Logit. *Econometrica*, 72, 1859-1876.

Manski, C.F. (1987). Semiparametric Analysis of Random Effects Linear Model from Binary Panel Data. *Econometrica*, 55, 357-362.

McFadden, D. (1974). Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Econometrics*. New York: Academic Press.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica* 46, 69-85.

National Credit Union Administration. (2011). *2011 Annual Report*.

Newey, W.K. (1984). A Method of Moments Interpretation of Sequential Estimators. *Economic Letters*, 14, 201-206.

Olsen, R.J. (1980). A Least Squares Correction for Selectivity Bias. *Econometrica*, 48 (7), 1815-1820.

Petersen, M.A., Rajan, R.G. (2002). Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance*, 57 (6), 2533-2570.

Ridder, G. (1990). Attrition in Multi-Wave Panel Data. In J. Hartog, G. Ridder, J. Theeuwes (Eds.), *Panel Data and Labor Market Studies*. Amsterdam: North-Holland Publishing.

Ridder, G. (1992). An Empirical Evaluation of Some Models for Non-Random Attrition in Panel Data. *Structural Change and Economic Dynamics*, 3, 337-355.

Rochina-Barrachina, M.E. (1999). A New Estimator for Panel Data Sample Selection Models. *Annales d'Économie et de Statistique*, 55/56, 153-181.

Semykina, A., Wooldridge, J.M. (2010). Estimating Panel Data Models in the Presence of Endogeneity and Selection. *Journal of Econometrics*, 157, 375-380.

Semykina, A., Wooldridge, J.M. (2011). Estimation of Dynamic Panel Data Models with Sample Selection. *Journal of Applied Econometrics*, 28, 47-61.

Sickles, R.C., Taubman, P. (1986). A Multivariate Error Components Analysis of the Health and Retirement Study of the Elderly. *Econometrica*, 54, 1339-1356.

Smith, D.J. (1984). A Theoretic Framework for the Analysis of Credit Union Decision Making. *Journal of Finance*, 39 (4), 1155-1168.

Trost, R.P., Lee, L.F. (1984). Technical Training and Earnings: A Polychotomous Choice Model with Selectivity. *Review of Economics and Statistics*, 66 (1), 151-156.

Vella, F., Verbeek, M. (1999). Two-step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias. *Journal of Econometrics*, 90, 239-263.

Verbeek, M. (1990). On the Estimation of a Fixed Effects Model with Selectivity Bias. *Economic Letters*, 34, 267-270.

Verbeek, M., Nijman, Th.E. (1996). Incomplete Panels and Selection Bias (Chapter 8). In L. Mátyás, P. Sevestre (Eds.). *The Econometrics of Panel Data: A Handbook of the Theory with Applications*. Dordrecht: Kluwer Academic Publishers.

Walter, J.R. (2006). Not Your Father's Credit Union. *Federal Reserve Bank of Richmond Economic Quarterly*, 92 (4), 353-377.

Wheelock, D.C., Wilson, P.W. (2011). Are Credit Unions Too Small? *Review of Economics and Statistics*, 93 (4), 1343-1359.

Wheelock, D.C., Wilson, P.W. (2013).The Evolution of Cost-Productivity and Efficiency among US Credit Unions. *Journal of Banking and Finance*, 37, 75-88.

Wilcox, J. (2005). Economies of Scale and Continuing Consolidation of Credit Unions. *Federal Reserve Banks of St. Francisco Economic Letter*, Number 2005-29.

Wilcox, J. (2006). Performance Divergence of Large and Small Credit Unions. *Federal Reserve Banks of St. Francisco Economic Letter*, Number 2006-19.

Wooldridge, J.M. (1995). Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions. *Journal of Econometrics*, 68, 115-132.

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

APPENDIX

TABLE A1. Call Report Definitions of the Variables

Variable	NCUA Account Definition	Description
y1	Acct_703 + Acct_386	Real estate loans: first mortgage real estate loans, other real estate loans
y2	Acct_475	Commercial loans: business and agricultural loans (MBLs) granted YTD
y3	Acct_025B - ***y1*** - ***y2***	Consumer loans: total loans, less real estate loans, less commercial loans
y4	Acct_799	Total investments
$\tilde{y}5$	(Acct_380 + Acct_381)/ Acct_018	Average interest rate on saving deposits: dividends on shares, interest on deposits, divided by total shares and deposits
$\tilde{y}6$	(Acct_110 + Acct_131)/ Acct_025B	Average interest rate on loans: total (gross) interest and fee income on loans, fee income, divided by total loan and leases
w1	(Acct_230 + Acct_250 + Acct_260 + Acct_270 + Acct_280 + Acct_290 + Acct_310 + Acct_320 + Acct_360)/Acct_018	Price of capital: travel and conference expense, office occupancy expense, office operations expense, educational and promotional expense, loan servicing expense, professional and outside services, member insurance, operating fees (examination and/or supervision fees), miscellaneous operating expenses, divided by total shares and deposits
w2	Acct_210/(Acct_564A + 0.5*Acct_564B)	Price of labor: employee compensation and benefits, divided by full-time equivalent employees [Number of credit union employees who are: Full-time (26 hours or more)+0.5*Part-time (25 hours or less per week)]
C	Acct_010	Total variable, noninterest cost: total non-interest expenses
$\tilde{k}$	Acct_931 + Acct_668 + Acct_945 + Acct_658 + Acct_940 + Acct_602	Equity: regular reserves, appropriation for non-conforming investments, accumulated unrealized gains (losses) on available-for-sale securities and other comprehensive income, other reserves, undivided earnings, net income
Total Assets	Acct_010	Total assets
Leverage	(Acct_860C + Aacct_820a + Acct_825 + Acct_018)/ Acct_010	Total liabilities [total borrowing, accrued dividends and interest payable on shares and deposits, accounts payable and other liabilities, total shares and deposits], divided by total assets
Reserves	Acct_931 + Acct_668	Regular reserves, appropriation for non-conforming investments
Current Members #	Acct_083	Total number of current members
Potential Members #	Acct_084	Total number of potential members

IN-TEXT TABLES AND FIGURES

TABLE 1. Zero-Value Observations, 1994-2011

Year	*y1*	*y2*	*y3*	*y4*	Total Obs.	Year	*y1*	*y2*	*y3*	*y4*	Total Obs.
1994	3,670	9,063	0	3	9,783	**2004**	2,344	7,099	1	64	8,209
1995	3,517	9,056	0	0	9,734	**2005**	2,171	6,695	1	57	7,948
1996	3,555	9,162	0	2	9,891	**2006**	2,044	6,333	1	68	7,718
1997	3,441	9,059	0	0	9,765	**2007**	1,952	6,101	1	59	7,506
1998	3,269	8,811	0	0	9,561	**2008**	1,805	5,703	1	38	7,174
1999	3,140	8,650	0	55	9,426	**2009**	1,485	5,086	1	55	6,521
2000	2,925	8,442	0	75	9,195	**2010**	1,612	5,306	1	115	6,761
2001	2,764	8,114	0	61	8,932	**2011**	1,539	5,212	1	61	6,591
2002	2,601	7,739	0	61	8,611	***Total***	*46,377*	*133,152*	*9*	*870*	*151,817*
2003	2,543	7,521	1	96	8,491						

NOTES: The variables are defined as follows. *y1* – real estate loans, *y2* – business and agricultural loans; *y3* – consumer loans; *y4* – investments.

TABLE 2. Tabulation of All Possible Heterogeneous Technologies

Technology	Obs.	Unique CUs	Technology	Obs.	Unique CUs
Complete Specialization			***Three-Output Specialization***		
y1	5	1	***y1, y2, y3***	20	10
y2	0	0	***y1, y2, y4***	0	0
y3	673	328	***y1, y3, y4***	87,122	11,764
y4	0	0	***y2, y3, y4***	526	306
Two-Output Specialization			***No Specialization***		
y1, y2	0	0	***y1, y2, y3, y4***	18,118	4,466
y1, y3	171	113			
y1, y4	4	1			
y2, y3	1	1			
y2, y4	0	0			
y3, y4	45,177	9,446			

NOTES: The variables are defined as follows. *y1* – real estate loans, *y2* – business and agricultural loans; *y3* – consumer loans; *y4* – investments.

TABLE 3. Summary Statistics, 1994-2011

Variable	Mean	Min	1st Qu.	Median	3rd Qu.	Max
			Technology 1			
Cost	171.8	0.7	47.6	101.2	205.3	9,866.0
y3	2,648.0	0.9	680.4	1,566.0	3,284.0	16,387.6
y4	1,547.0	0.0	167.9	580.3	1,635.0	262,500.0
$\tilde{y}5$	0.028	0.000	0.017	0.029	0.038	0.056
$\tilde{y}6$	0.100	0.000	0.082	0.095	0.110	0.993
w1	0.026	0.000	0.016	0.023	0.031	0.695
w2	32.9	0.0	20.1	32.2	43.3	266.3
$\tilde{k}$	687.6	0.6	175.9	386.7	826.0	54,030.0
Total Assets	4,712.0	22.3	1,215.0	2,769.0	5721.0	373,600.0
Leverage	0.009	0.000	0.002	0.004	0.010	0.842
Reserves	198.8	0.0	47.6	100.2	214.0	18,270.0
Current Members #	1,127	27	401	745	1,378	43,560
Potential Members #	4,389	1	700	1461	3,000	10,000,000
Multiple Bond CU	0.321					
Federal CU	0.625					
State CU (insured)	0.360					
			Technology 2			
Cost	2,244.0	3.2	333.4	767.5	1,965.0	580,500.0
y1	15,780.0	0.0	675.0	2,850.0	10,290.0	6,501,000.0
y3	24,750.0	3.0	3,767.0	8,172.0	20,090.0	9,126,000.0
y4	18,290.0	0.0	1,683.0	4,859.0	13,300.0	4,620,000.0
$\tilde{y}5$	0.026	0.000	0.016	0.027	0.036	0.194
$\tilde{y}6$	0.091	0.000	0.079	0.089	0.100	0.973
w1	0.026	0.000	0.016	0.023	0.031	0.695
w2	46.6	0.0	37.8	45.2	54.1	6,187.0
$\tilde{k}$	7,338.0	0.8	1,080.0	2,477.0	5,955.0	2,587,000.0
Total Assets	65,750.0	116.0	8,908.0	20,580.0	51,300.0	24,090,000.0
Leverage	0.010	0.000	0.002	0.005	0.010	0.351
Reserves	2,638.0	0.0	294.7	707.5	1,800.0	2,563,000.0
Current Members #	8,859	5	1,754	3,570	8,276	2,451,000
Potential Members #	72,790	1	3,500	9,000	32,430	27,000,000
Multiple Bond CU	0.427					
Federal CU	0.610					
State CU (insured)	0.378					

TABLE 3. (cont.)

Variable	Mean	Min	1st Qu.	Median	3rd Qu.	Max
			Technology 3			
Cost	10,030.0	18.3	1,306.0	3,619.0	10,230.0	1,448,000.0
y1	119,400.0	1.0	8,314.0	29,230.0	94,810.0	18,940,000.0
y2	5,831.0	0.0	163.7	710.9	3,577.0	874,500.0
y3	98,490.0	13.0	10,260.0	29,440.0	84,190.0	14,340,000.0
y4	66,820.0	3.0	4,599.0	14,620.0	48,050.0	12,360,000.0
$\tilde{y}1$	0.024	0.000	0.015	0.023	0.033	0.067
$\tilde{y}2$	0.083	0.000	0.072	0.082	0.093	0.873
w1	0.026	0.000	0.016	0.023	0.031	0.695
w2	51.6	0.2	42.2	49.9	58.7	324.4
$\tilde{k}$	32,970.0	10.0	3,902.0	10,250.0	29,870.0	5,079,000.0
Total Assets	326,400.0	224.0	35,860.0	98,320.0	288,600.0	46,930,000.0
Leverage	0.023	0.000	0.004	0.009	0.021	0.439
Reserves	11,880.0	0.0	1,106.0	2,956.0	8,159.0	4,906,000.0
Current Members #	32,070	119	4,972	12,570	33,070	3,867,000
Potential Members #	365,800	250	15,000	66,500	250,000	28,000,000
Multiple Bond CU	0.307					
Federal CU	0.523					
State CU (insured)	0.457					

NOTES: The variables are defined as follows. *Cost* – total variable, non-interest cost; *y1* – real estate loans, *y2* – business and agricultural loans; *y3* – consumer loans; *y4* – investments; $\tilde{y}5$ - average saving pricing; $\tilde{y}6$ – average loan pricing; *w1* – price of capital; *w2* – price of labor; $\tilde{k}$ – equity capital; *Leverage* – the ratio of total debt to total assets; *Multiple Bond*, *Federal*, and *State (insured)* CU – indicator variables that take value of one if a CU is multiple-bond, federally accredited, or state-accredited (but federally insured), respectively. The remaining variables are self-descriptive. *Cost, y1, y2, y3, y4, w2,* $\tilde{k}$, *Assets, Reserves* are in thousands of real 2011 US dollars; $\tilde{y}5$, $\tilde{y}6$, *w1, Leverage* are interest rates and thus are unit-free. The numbers of *Current* and *Potential Members* are in terms of number of people. Despite that minima of several variables are reported to be zeros (due to rounding), they are not exactly equal to zeros.

TABLE 4. Summary of Returns to Scale Estimates, Models 1 through 3

	Point Estimates of RS							Categories of RS			
Model	Mean	St. Dev.	Min	1st Q	Median	3rd Q	Max	DRS	CRS	IRS	Total
				Technology 1							
(1)	1.780	0.247	1.095	1.607	1.744	1.918	3.048	0	1	44273	44274
(2)	1.217	0.133	0.845	1.132	1.197	1.277	3.143	359	471	43444	44274
(3)	1.156	0.072	0.886	1.107	1.144	1.192	1.989	35	56	44183	44274
				Technology 2							
(1)	1.115	0.105	0.801	1.040	1.111	1.184	1.636	10626	2626	72129	85381
(2)	1.076	0.053	0.868	1.038	1.074	1.111	1.481	3981	3041	78359	85381
(3)	1.080	0.059	0.877	1.037	1.074	1.115	1.511	4535	2132	78714	85381
				Technology 3							
(1)	1.075	0.041	0.879	1.053	1.071	1.090	1.804	13	215	17529	17757
(2)	1.059	0.040	0.866	1.037	1.055	1.074	1.781	29	659	17069	17757
(3)	1.035	0.053	0.889	0.999	1.026	1.061	1.438	4081	1153	12523	17757
				Whole Sample							
(1)	1.310	0.347	0.801	1.067	1.155	1.554	3.048	10639	2842	133931	147412
(2)	1.117	0.107	0.845	1.049	1.089	1.153	3.143	4369	4171	138872	147412
(3)	1.097	0.075	0.877	1.044	1.089	1.140	1.989	8651	3341	135420	147412

TABLE 5. Summary of Returns to Scale Estimates, Models 4 through 6

	Point Estimates of RS							Categories of RS			
Model	Mean	St. Dev.	Min	1st Q	Median	3rd Q	Max	DRS	CRS	IRS	Total
				Technology 1							
(4)	1.921	0.364	1.033	1.664	1.845	2.103	3.609	0	1	44273	44274
(5)	1.252	0.191	0.874	1.160	1.227	1.312	3.617	128	272	43874	44274
(6)	1.195	0.102	0.922	1.132	1.178	1.239	3.776	11	28	44235	44274
				Technology 2							
(4)	1.137	0.095	0.857	1.070	1.134	1.200	1.589	4836	2133	78412	85381
(5)	1.107	0.060	0.922	1.063	1.102	1.145	1.510	911	1397	83073	85381
(6)	1.106	0.071	0.905	1.055	1.097	1.146	1.628	2335	1306	81740	85381
				Technology 3							
(4)	1.209	0.064	1.056	1.165	1.208	1.244	1.944	0	2	17755	17757
(5)	1.102	0.041	0.959	1.078	1.101	1.120	1.755	0	23	17734	17757
(6)	1.056	0.060	0.895	1.014	1.047	1.086	1.457	2206	1006	14545	17757
				Whole Sample							
(4)	1.381	0.414	0.857	1.115	1.204	1.598	3.609	4836	2136	140440	147412
(5)	1.150	0.133	0.874	1.079	1.122	1.189	3.617	1039	1692	144681	147412
(6)	1.127	0.093	0.895	1.063	1.115	1.174	3.776	4552	2340	140520	147412

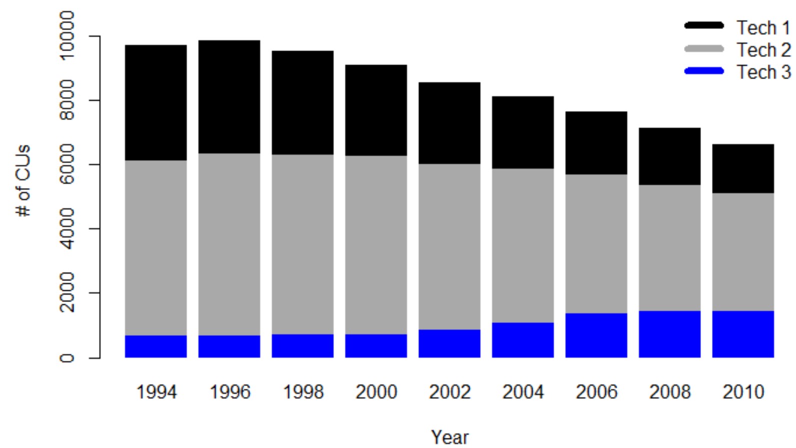

FIGURE 1. Tabulation of Credit Unions
by Technology Type

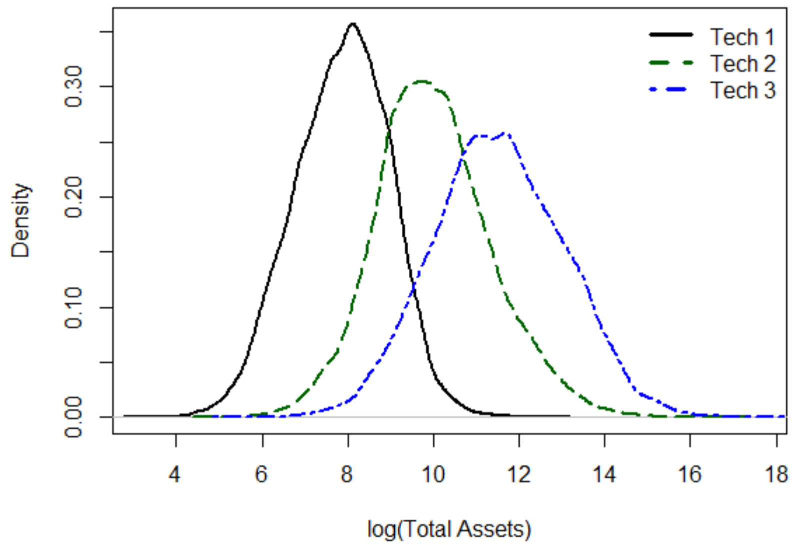

FIGURE 2. Kernel Densities of (log) Total Assets
Tabulated by Technology Type, 1994-2011

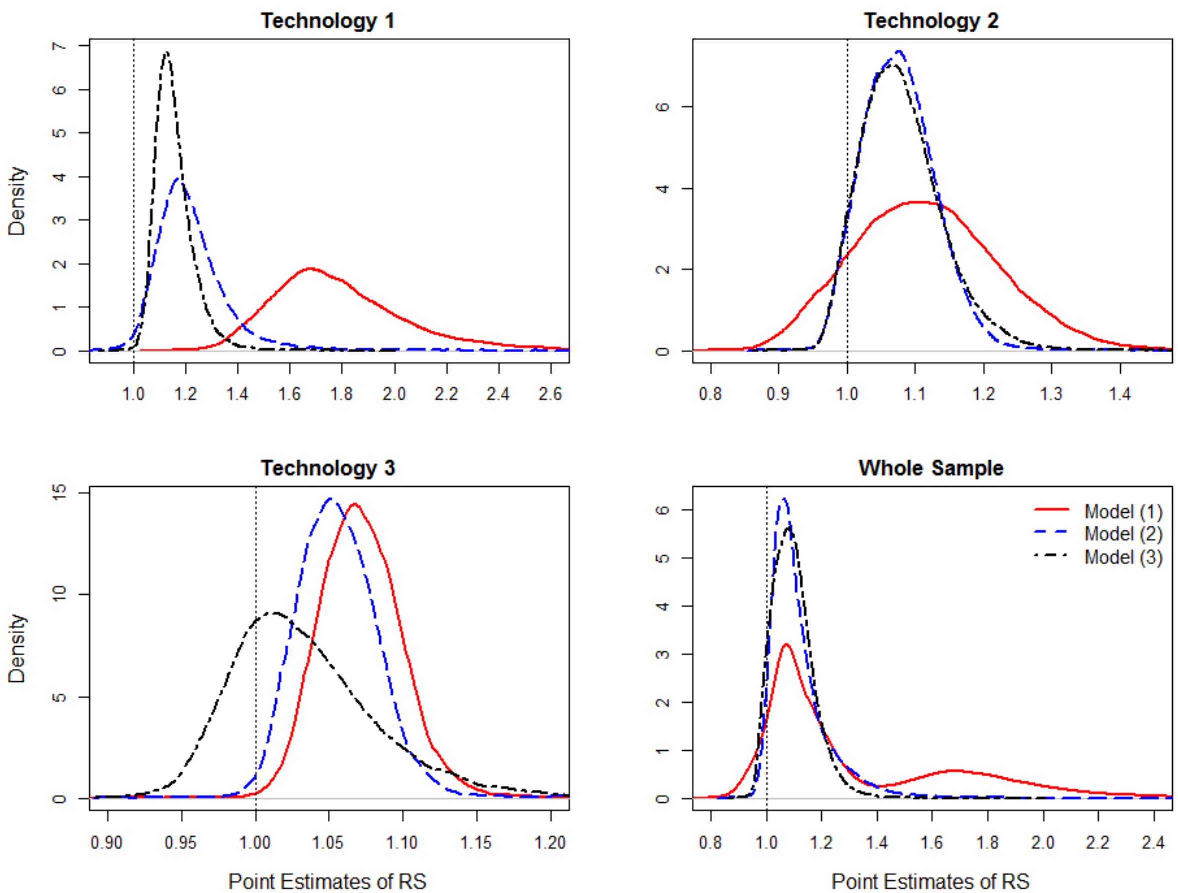

FIGURE 3. Kernel Densities of Returns to Scale Estimates from Models 1 through 3.

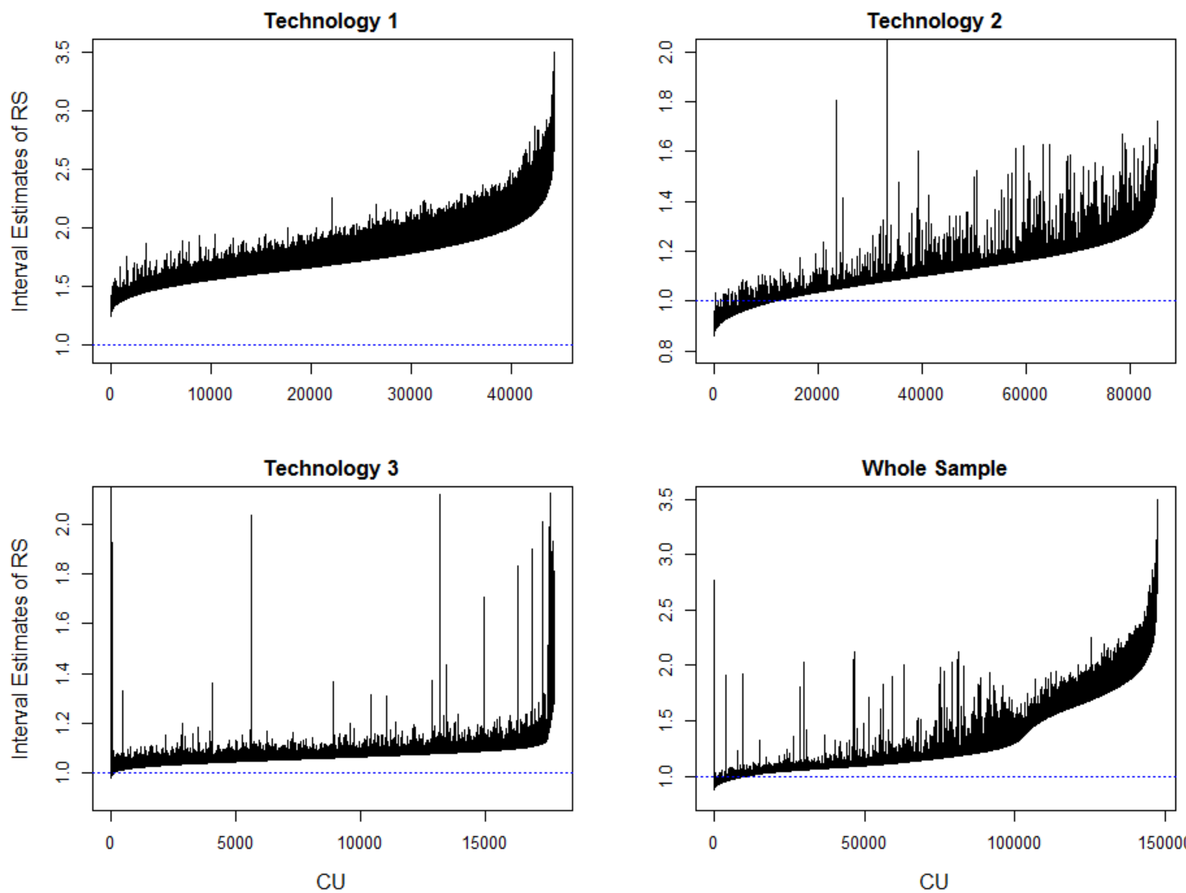

FIGURE 4. The 95% Confidence Intervals of Returns to Scale Estimates from Model 1.

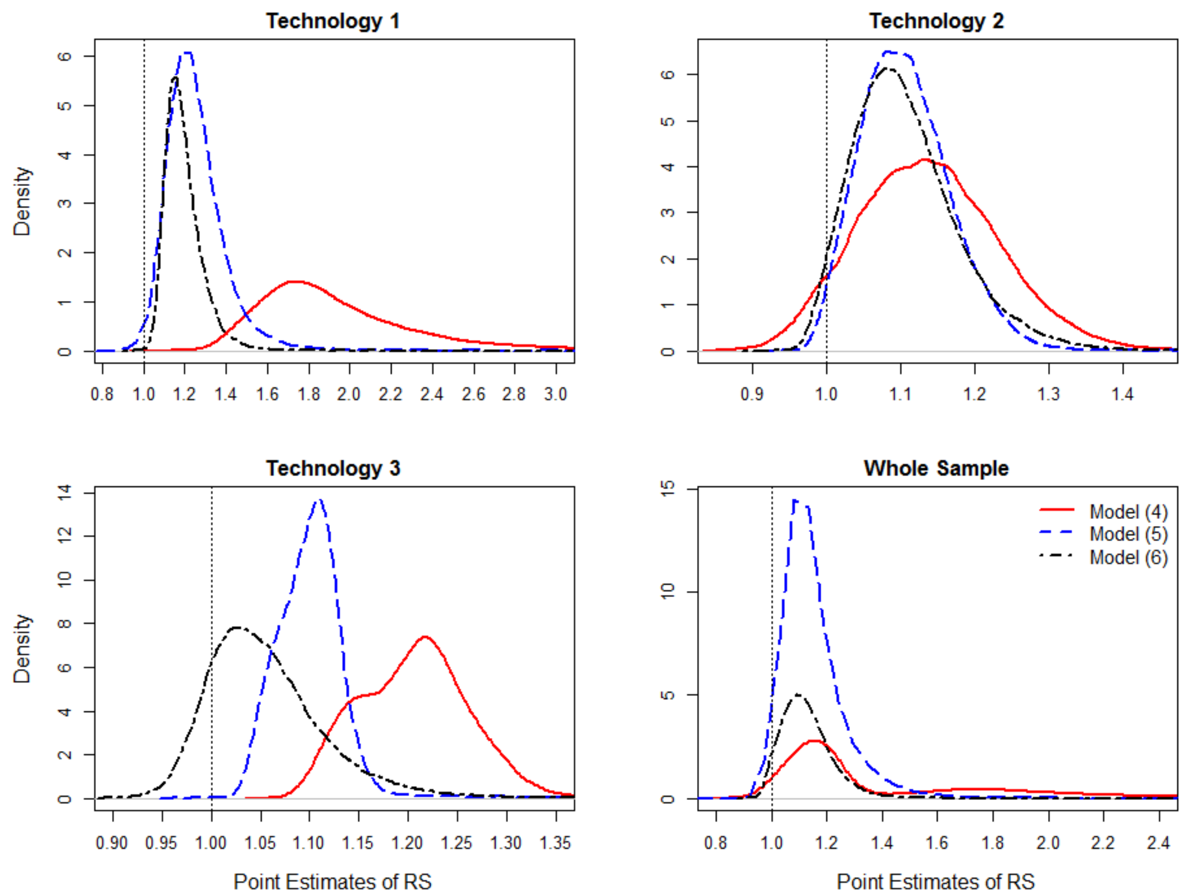

FIGURE 5. Kernel Densities of Returns to Scale Estimates from Models 4 through 6.

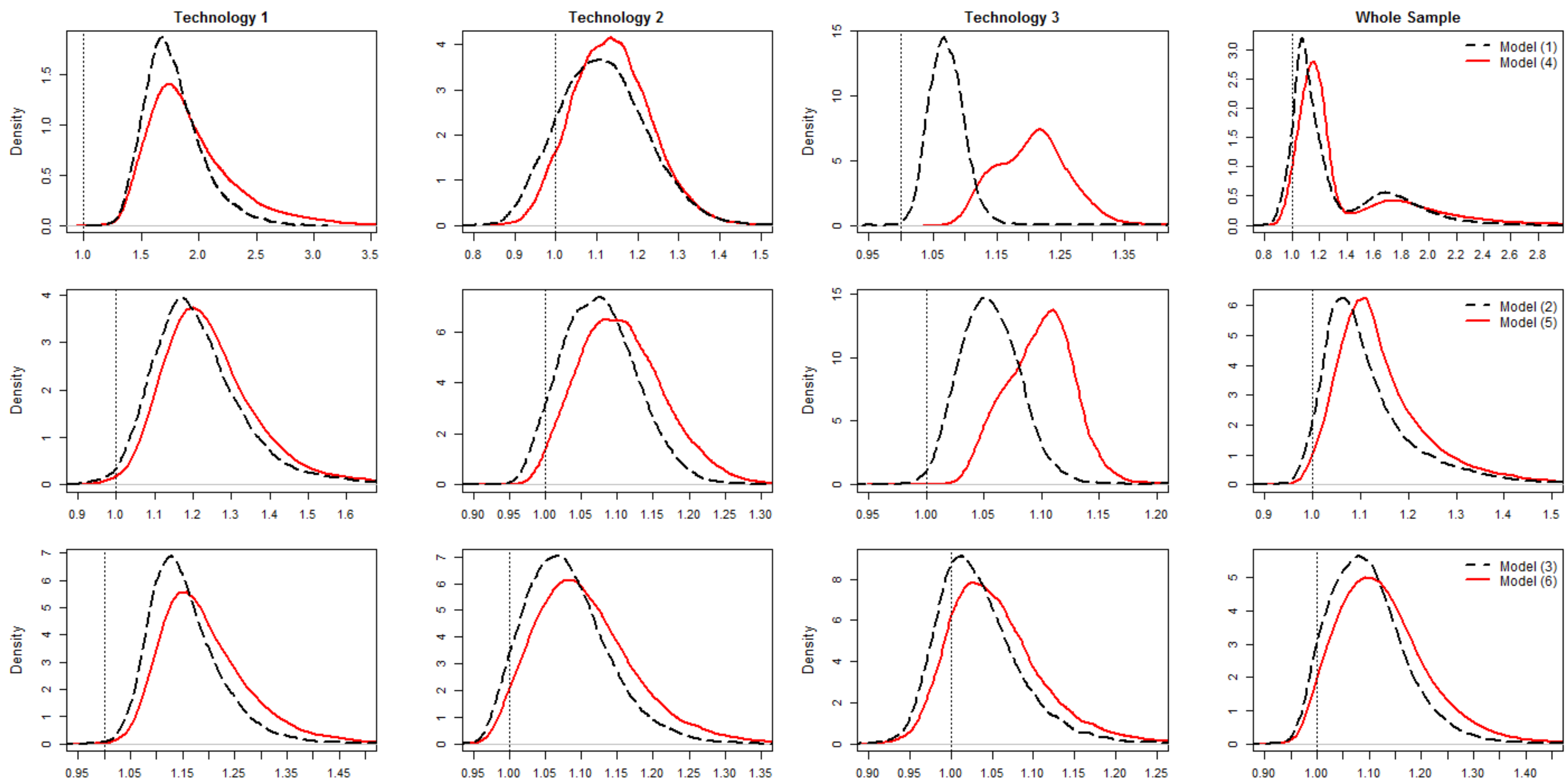

FIGURE 6. Kernel Densities of Returns to Scale Estimates from Models 1 through 6.

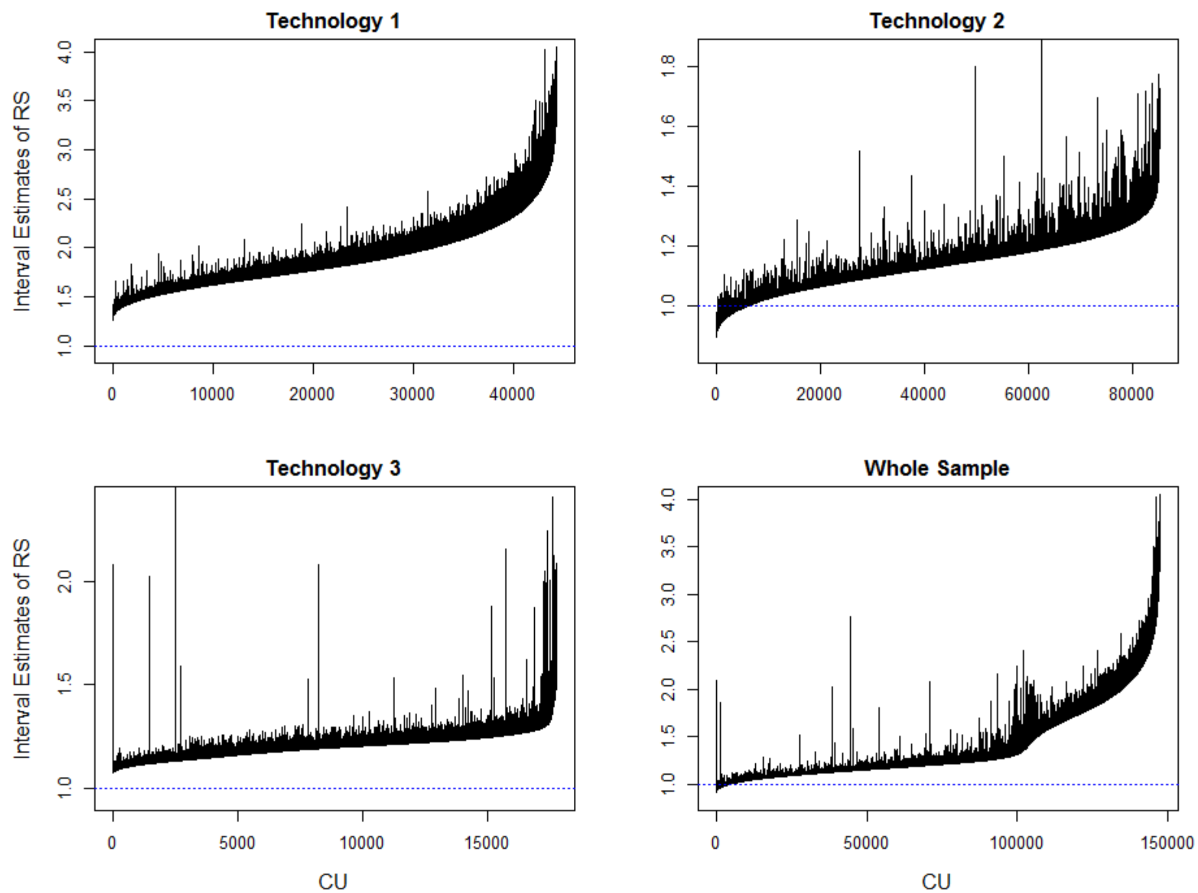

FIGURE 7. The 95% Confidence Intervals of Returns to Scale Estimates from Model 4.

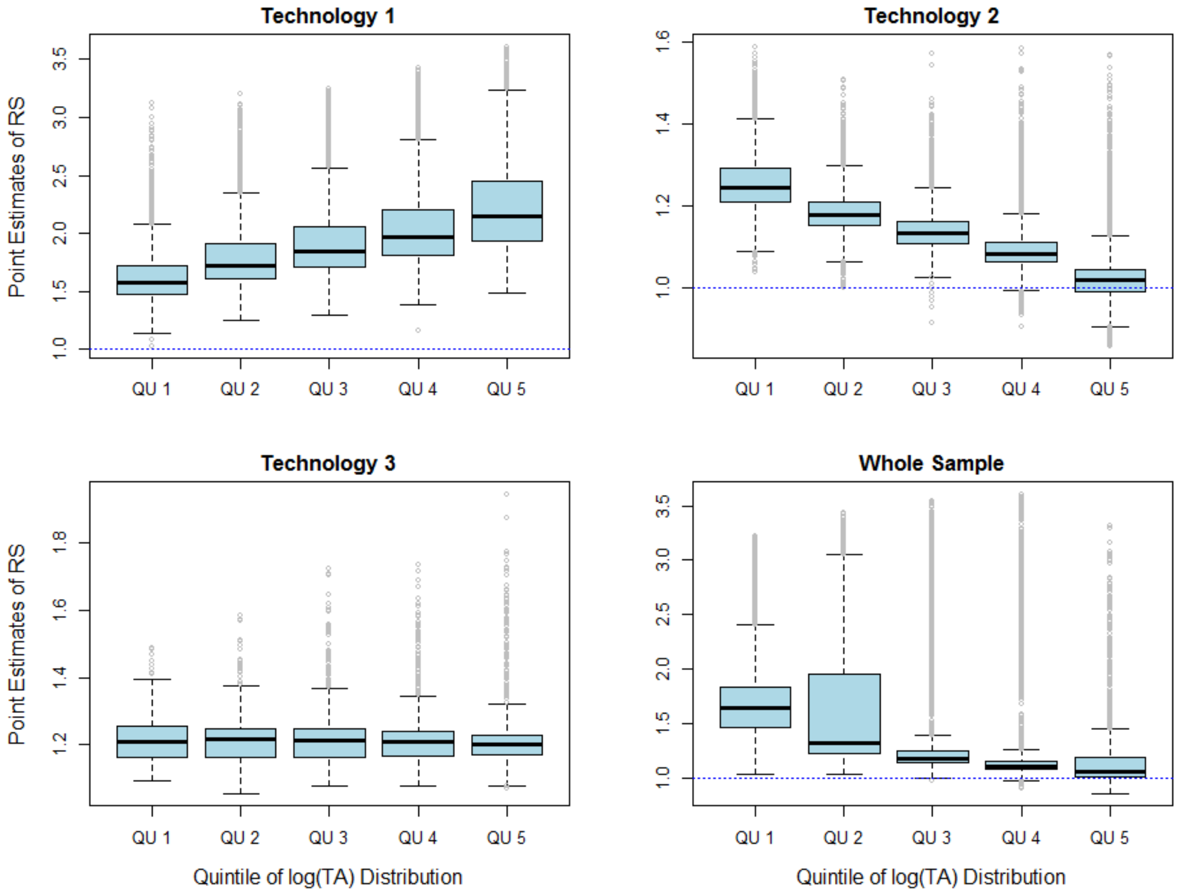

FIGURE 8. Returns to Scale by (log) Total Assets Quintiles;
Estimates from Model 4.