Inefficiency and Bank Failures: A Joint Bayesian Estimation of a Stochastic Frontier Model and a Hazards Model

Jim Sánchez González[*]
Diego Restrepo Tobón[†]
Andrés Ramírez Hassan[‡]

School of Economics and Finance
Universidad EAFIT, Medellín, Colombia

October 24, 2018

Abstract

Estimating inefficiency separately from the hazards model results in inefficient, biased, and inconsistent estimators. We develop a method to simultaneously estimate a stochastic frontier model and a hazards model using Bayesian techniques. This method overcomes issues related to two-stage estimation methods, allows for computing the marginal effects of the inefficiency over the probability of failure, and facilitates statistical inference of the functions of the parameters. Simulation exercises show that our method performs better than two-stage MLE. We find that inefficiency plays a statistically and economically significant role in determining the time to failure of U.S. commercial banks.

Key words: Bank Failures; Proportional Hazards Model; Technical Inefficiency
JEL Classification: C11; G21; G33

---

[*]Corresponding author's email: jsanch83@eafit.edu.co
[†]Email: drestr16@eafit.edu.co
[‡]Email: aramir21@eafit.edu.co

1 Introduction

Literature regarding the estimation of the probability of failure for a given firm considers different explanatory variables like financial ratios (the ones known as CAMELs[1] for instance) and economic variables (unemployment, inflation, Gross Domestic Product). More recently, some articles include inefficiency–as proxy for managerial practices–as a possible explanatory variable for the probability of failure. Given the unobservable characteristic of inefficiency, researchers usually opt for a two-stage approach (see for instance, Barr, Seiford, and Siems (1993); A. N. Berger and Humphrey (1992)). That is, estimating the inefficiency in a first step and then use these estimates as a covariate for the second equation.

Such methodology yields inefficient, inconsistent, and biased estimators. This is because the second equation does not consider the effect of the measurement error of the first estimation methods. Using Bayesian techniques, we propose a method for the simultaneous estimation of inefficiencies and time to failure using a stochastic frontier model and a proportional hazards model. Doing so, we avoid the econometric flaws associated with two-stage estimation. Besides, our Bayesian approach allows performing statistical inference of functions of the parameters such as elasticities, returns to scale, and individual inefficiencies without any extra mathematical or computational burden Ramirez Hassan and Montoya Blandon (2017).

Understanding why some firms fail while others thrive is an important issue for economic policy and firms' management. This is even more important if the failing firm can cause massive economic distress as banks do. Consequently, several studies attempt to explain the causal relationship between different economic and financial variables and the probability of failure.[2] Within the different explanatory variables considered in the literature, inefficiency is of particular interest. In general and economically talking, inefficiency is an undesirable characteristic because resources are scarce. Besides, less inefficient societies serve better their inhabitants needs. In other words, more inefficient firms have either higher cost or lower production levels which, under an imperfect market structure, impact negatively the final consumers.

Including an inefficiency measure as covariate has the inconvenience that it is not observable and, consequently, it must be estimated. Most of the literature takes a two-stage approach to tackle this problem; e.g. the researchers use a model, like data envelope analysis or stochastic frontier, estimate the inefficiency, and use the result as an explanatory variable for either the hazards, the probit, or the logit models. However, this approach has some econometric flaws; namely, that the estimates of the second model are inefficient,

---

[1]CAMELs stands for Capital adequacy, Assets quality, Managerial practices, Earnings quality, and Liquidity position.

[2]See for instance: Altman (1968); Beaver (1966); Brewer III, Genay, Hunter, and Kaufman (2003); Kumar and Ravi (2007); Ohlson (1980). These studies do not explicitly estimate the effect of inefficiency on the probability of failure, but they do find a correlation between the more inefficient firms and the ones that failed.

biased, and inconsistent because the second step does not consider the estimation error associated with the first model.

Up to the best of our knowledge, there is just one article (Tsionas & Papadogonas, 2006) and one working paper (Almanidis & Sickles, 2016) that estimate inefficiency and the probability of failure simultaneously. Tsionas and Papadogonas (2006) use Gaussian Quadrature and maximum likelihood for estimating a probit, a logit, and an extreme value model including technical inefficiency as a covariate for manufacturing firms in Greece. On the other hand, Almanidis and Sickles (2016) use simulated maximum likelihood to estimate a mixture hazards model including a stochastic frontier model for the inefficiency.

Our contribution to the literature is twofold. First, we propose an estimation method that solves the econometric flaws seen in literature. By estimating simultaneously the inefficiency and the time to failure, one can reduce the estimation error of the time to failure. Hence, the method we propose considers the simultaneous estimation of a stochastic frontier model to estimate the inefficiency and a proportional hazards model to estimate the time to failure. This simultaneous estimation is developed within a Bayesian framework. The reason for using Bayesian techniques is that it allows obtaining the conditional posterior distribution of the parameters which can be readily used to perform statistical inference of the functions of the parameters. Additionally, using Bayesian techniques allow for obtaining the whole empirical distribution of the inefficiency for each firm, instead of just a point estimate of such inefficiency per firm (see Koop (2003)).

This article differentiates from that of Tsionas and Papadogonas (2006) in two aspects. First, we use a proportional hazards model instead of classification models, namely the probit and logit models. Proportional hazards models are more flexible in the sense that they allow for censored data[3] and they model the mean time to failure instead of just binary responses, the firm will or will not fail in the next period. Second, we use commercial banks data for the empirical application while Tsionas and Papadogonas use data of Greek manufacturing firms.

Likewise, our research differentiates from that of Almanidis and Sickles (2016) because our aim is to understand the effects of inefficiency over the mean time to failure and to perform statistical inference of the functions of the parameters estimates such as elasticities and returns to scale. Though Almanidis and Sickles also use a simultaneous estimation method based on the expectation-maximization algorithm, the authors' goal is to develop an early warning model to determine if more inefficient banks should receive *on-site* risk assessment by the Federal Deposit Insurance Corporation (FDIC). Additionally, our study can be regarded, in some sense, complementary to this one because we analyze a wider time lapse, from 2001 to 2010, while Almanidis and Sickles just use data for 2008-2009.

This article is enclosed in two econometric frameworks: proportional hazards models

---

[3]Censored data refers to observations for which the time recorded is not a failure time but the time observed at the end of the sampling period. In other words, at the end of the sample period, the individual under consideration was still alive.

and stochastic frontier models. There are plenty of literature regarding survival analysis,[4] most of it in Medicine and Biology, though also in Finance and Economics.[5] In our case, the aim is to model the banks' time to failure given a set of covariates. Proportional hazards models, proposed by Cox (1972), are very common in the literature of survival analysis because their practicality, easy estimation, and because they i) allow for modeling not just the probability of failure but also the time to failure; ii) they allow considering censored data (models like discriminant analysis and the probit or logit models do not); and iii) they do not make strong assumptions about the distribution of the duration times–one can assume Exponential, Weibull, or Gamma distributions depending on one particular needs.[6]

Regarding stochastic frontier models, there is also substantial literature. Since the initial works of Aigner, Lovell, and Schmidt (1977); Battese and Corra (1977); Meeusen and van Den Broeck (1977) several authors have used stochastic frontier models to estimate technical inefficiency and allocative inefficiency.[7] The main difference between these works is the assumptions made over the distribution of the inefficiency term.[8]

Despite the abundant literature regarding these models; few authors use them jointly to understand the effects of inefficiency on the probability of failure; though the correlation between them is acknowledged since the early 90's. For instance, Barr et al. (1993); A. N. Berger and Humphrey (1992) do not explicitly estimate the impact of inefficiency on failure, but they do find a positive correlation between inefficiency and the probability of failure; i.e., more inefficient banks are more likely to fail than efficient ones. Similarly, Assaf, Berger, Roman, and Tsionas (2017) find that more cost-efficient banks, in normal times, have less probability to fail in the subsequent crisis.

Up to our knowledge, just Barr and Siems (1997); Dimara, Skuras, Tsekouras, and Tzelepis (2008); Wheelock and Wilson (1995, 2000) include inefficiency measures as an explanatory variable for the hazards model. However, they estimate their models in a two-stage fashion; that is, the authors estimate either technical or allocative inefficiency using data envelope analysis (DEA) or stochastic frontier models and then include the estimated inefficiency into the estimation of the hazards model.

The rest of the article is organized as follows: section 2 outlines the Bayesian modeling strategy for the stochastic frontier model and the hazards model and presents some simulations results. Section 3 presents the details of the data as well as the models specification and the variables construction for the empirical application which is presented in section 4. Finally, in section 5 we conclude.

---

[4]Hazards models are also known in the literature as Survival models or Duration models.

[5]See for instance: Cole and Gunther (1998); Halling and Hayden (2006); Tsionas and Papadogonas (2006).

[6]For applications of proportional hazards model see for instance: Cole and Wu (2009); Shumway (2001); Whalen (1991); Wheelock and Wilson (1995, 2000).

[7]There are other methodologies for estimating inefficiency; a common one in the literature is for instance data envelope analysis (DEA).

[8]For an extensive review of stochastic frontier models and their estimation issues see Kumbhakar and Lovell (2000); Kumbhakar, Wang, and Horncastle (2015).

4

2   Methodology

Since the two-stage approach is not suitable for the reasons mentioned before, we develop a method based on Bayesian techniques for estimating both models simultaneously. The advantages of using a Bayesian approach is that we can treat the unobserved inefficiencies as parameters. Hence, we can use data augmentation techniques (see Tanner and Wong (1987)) to obtain the posterior distribution of the inefficiencies. In addition, we can easily perform statistical inference of functions of the parameter estimates.

In the coming subsections, we provide–very briefly–the fundamental equations on which the Bayesian theory is based–that is the *Bayes theorem*. Then, we present the definition of the priors to be used for each parameter and derive the likelihood functions for the stochastic frontier model and the proportional hazards model. Finally, we present the analytical form of the conditional posterior distribution for each parameter.

Bayes theorem states that the posterior distribution of the parameters given the data is proportional to the product of the likelihood function times the priors of the parameters, algebraically:

$$\pi\left(\theta|\text{Data}\right) = \frac{\mathcal{L}\left(\text{Data}|\theta\right)\pi\left(\theta\right)}{\int_{\Theta}\mathcal{L}\left(\text{Data}|\theta\right)\pi\left(\theta\right)d\theta} \qquad (2.1)$$
$$\propto \mathcal{L}\left(\text{Data}|\theta\right)\pi\left(\theta\right)$$

where the left-hand-side term is the posterior distribution of the parameters (θ). The term in the left of the right-hand-side of the equation is the likelihood function of the data given the parameters, and the term in the right is known as the prior of the parameters.

2.1   Likelihood Functions

2.1.1   Stochastic Frontier Model

Consider the cost function for the stochastic frontier model (see Kumbhakar et al. (2015, p. 117)),

$$\ln C_i^a = \ln C_i^*\left(w,\,y\right) + \mu_i + \eta_i \qquad (2.2)$$

where: C_i^a is the actual cost of each firm and C_i^* is the efficient cost function. As it is common in the literature we assume that μ is a two-sided normally distributed error term with mean zero and error precision h_μ.[9] Likewise, we assume that the inefficiency term (η) distributes half-normal with mean zero and error precision h_η.[10] That is,

---

[9]The error precision is the inverse of the variance, $\sigma^2 = \frac{1}{h}$.

[10]Common assumptions for the inefficiency term also include the exponential distribution and the truncated normal distribution. However, as reported in Kumbhakar et al. (2015, p. 94) different distributional assumptions do not affect the results considerably.

$$\mu_i \sim \mathcal{N}\left(0, h_\mu^{-1}\right) \qquad \text{and} \qquad \eta_i \sim \mathcal{N}^+\left(0, h_\eta^{-1}\right) \qquad (2.3)$$

Given the above assumptions, definitions, and expressing $\ln C_i^a = c_i$ and $\ln C_i^*(w, w) = x_i'\beta$; we have that the likelihood function for the stochastic frontier model given the inefficiencies can be represented as,

$$
\begin{aligned}
\mathcal{L}\left(c \,|\, \beta, h_\mu, \eta, x\right) &= \prod_i \frac{h_\mu^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}} \exp\left[-\frac{1}{2}h_\mu\left(c_i - x_i'\beta - \eta_i\right)'\left(c_i - x_i'\beta - \eta_i\right)\right] \\
&\propto h_\mu^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\mu \sum_i \left(c_i - x_i'\beta - \eta_i\right)'\left(c_i - x_i'\beta - \eta_i\right)\right]
\end{aligned}
\qquad (2.4)
$$

2.1.2 Proportional Hazards Model

Now, let us consider the proportional hazards model. Let y_i be the duration time for the ith bank and let ν_i be a censoring indicator; where $\nu_i = 1$ if y_i is a failure time and $\nu_i = 0$ if y_i is right censored. Then, the likelihood function for the proportional hazards model can be expressed as the product of the probability of failing times the probability of surviving:

$$\mathcal{L}\left(y|\gamma, \delta, \eta, z\right) = \prod_i f\left(z_i'\gamma + \delta\eta_i\right)^{\nu_i} \left[1 - F\left(z_i'\gamma + \delta\eta_i\right)\right]^{1-\nu_i} \qquad (2.5)$$

where $f\left(\cdot\right)$ is the probability distribution function (p.d.f.) and $F\left(\cdot\right)$ is the cumulative distribution function (c.d.f.) of the hazards model.

We assume an exponential distribution for the failure time, $y_i \sim \text{Exp}\left(\lambda_i\right)$. The reason to consider this probability density function is that it embraces proportional hazards models as well as accelerated failure time models. Besides, this p.d.f. is easy to interpret (see Cameron and Trivedi (2005)). To account for the covariates we assume that $\lambda_i = \exp\left(z_i'\gamma + \delta\eta_i\right)$. Then, the likelihood function of the proportional hazards model can be expressed as:

$$
\begin{aligned}
\mathcal{L}\left(y|\gamma, \delta, \eta, z\right) &= \prod_i \left\{\exp\left(z_i'\gamma + \delta\eta_i\right)\exp\left[-y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]\right\}^{\nu_i} \\
&\qquad\qquad\qquad\qquad\qquad \left\{\exp\left[-y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]\right\}^{1-\nu_i} \\
&= \exp\left[\sum_i \nu_i\left(z_i'\gamma + \delta\eta_i\right)\right]\exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]
\end{aligned}
\qquad (2.6)
$$

2.2 Conditional Posterior Distributions

From a Bayesian perspective, we can handle each firms inefficiency as unknown parameters. These, in turn, are drawn from a prior distribution (half normal) with hyperparameter h_η

which also have a prior distribution. That is, we are using a hierarchical structure in our model.

Defining $\theta = (\beta, \gamma, \delta, h_\mu, h_\eta)$ and applying the Bayes theorem, we can express the posterior distribution of the parameters given the data as:

$$
\begin{aligned}
\pi\left(\theta \mid y,\, c\right) &\propto \mathcal{L}\left(y,\, c \mid \theta,\, \eta\right) \cdot \pi\left(\theta,\, \eta\right) \\
&= \mathcal{L}\left(y|c,\, \theta,\, \eta\right) \cdot \mathcal{L}\left(c|\theta,\, \eta\right) \cdot \pi\left(\eta|\theta\right) \cdot \pi\left(\theta\right)
\end{aligned}
\tag{2.7}
$$

Taking into account that y and c are independent after conditioning on the inefficiency term, we have that:

$$
\pi\left(\theta|y,\, c\right) \propto \mathcal{L}\left(y|\theta,\, \eta\right) \cdot \mathcal{L}\left(c|\theta,\, \eta\right) \cdot \pi\left(\eta|\theta\right) \cdot \pi\left(\theta\right)
\tag{2.8}
$$

The first term on the right-hand side of equation 2.8 is the likelihood function of the proportional hazards model (equation 2.6), the next one is the likelihood function of the stochastic frontier model (equation 2.4), and $\pi\left(\eta|\theta\right)$ and $\pi\left(\theta\right)$ are the prior of the inefficiency given the parameters and the priors of the parameters.

2.2.1   Priors

The prior of the inefficiency term conditioned to the parameters is:

$$
\begin{aligned}
\pi\left(\eta|\theta\right) &= \prod_i \frac{2^{\frac{1}{2}} h_\eta^{\frac{1}{2}}}{\pi^{\frac{1}{2}}} \exp\left[-\frac{1}{2} h_\eta \eta_i^2\right] \\
&\propto h_\eta^{\frac{N}{2}} \exp\left[-\frac{1}{2} h_\eta \sum_i \eta_i^2\right]
\end{aligned}
\tag{2.9}
$$

Following the literature, we assume Normal priors for β, γ, and δ. That is,

$$
\beta \sim \mathcal{N}_k\left(B_0,\, \Sigma_0\right) \qquad \gamma \sim \mathcal{N}_j\left(\Gamma_0,\, \Omega_0\right) \qquad \delta \sim \mathcal{N}\left(\Delta_0,\, h_\delta^{-1}\right)
\tag{2.10}
$$

and assume Gamma distributions for the error precision parameters h_μ and h_η,

$$
h_\mu \sim \mathcal{G}\left(\frac{s_\mu}{2},\, \frac{v_\mu}{2}\right) \qquad\qquad h_\eta \sim \mathcal{G}\left(\frac{s_\eta}{2},\, \frac{v_\eta}{2}\right)
\tag{2.11}
$$

2.2.2   Conditional Posterior Distributions

It can be shown (see Appendix A.1) that after some algebra and a procedure known as completing the squares, the conditional posterior distribution of β is a multivariate Normal distribution. That is, $\beta|y,\, c \sim \mathcal{N}\left(\mu_1,\, B_1\right)$; where

$$B_1 = \left[h_\mu \sum_i x_i' x_i + \Sigma_0^{-1} \right]^{-1} \qquad \text{and} \qquad \mu_1 = B_1 \left[h_\mu \sum_i x_i' (c_i - \eta_i) + \Sigma_0^{-1} B_0 \right] \quad (2.12)$$

Likewise, it can be shown (see Appendix A.2 and A.3) that both error precision parameters distribute Gamma. As follows, $h_\mu | y,\, c \sim \mathcal{G}\left(\frac{s_{1\mu}}{2},\, \frac{v_{1\mu}}{2} \right)$; where

$$s_{1\mu} = s_\mu + N \qquad\qquad v_{1\mu} = \sum_i (c_i - x_i\beta - \eta_i)'(c_i - x_i\beta - \eta_i) + v_\mu \qquad\qquad (2.13)$$

and, $h_\eta \,|\, y,\, c \sim \mathcal{G}\left(\frac{s_{1\eta}}{2},\, \frac{v_{1\eta}}{2} \right)$; with,

$$s_{1\eta} = s_\eta + N \qquad\qquad v_{1\eta} = \sum_i \eta_i^2 + v_\eta \qquad\qquad (2.14)$$

Given the assumed priors and the likelihood functions for γ and δ, we cannot derive standard forms (as we do for β, h_μ, and h_η) for the conditional posterior distribution of γ and δ. However, we do have an analytical form (see Appendix A.4 and A.5 for their derivation) for such conditional posterior distributions.

$$\pi(\gamma | y,\, c) \propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i \right) - \sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i \right) - \frac{1}{2}(\gamma - \Gamma_0)' \Omega_0^{-1} (\gamma - \Gamma_0) \right] \quad (2.15)$$

$$\pi(\delta | y,\, c) \propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i \right) - \sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i \right) - \frac{1}{2}h_\delta (\delta - \Delta_0)' (\delta - \Delta_0) \right] \quad (2.16)$$

Finally, we can obtain the conditional posterior distribution for each firm's inefficiency. See Appendix A.6 for its derivation.

$$\pi(\eta \,|\, y\,,c) \propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i \right) - \sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i \right) \right.$$
$$\left. - \frac{1}{2}h_\mu \sum_i (c_i - x_i\beta - \eta_i)'(c_i - x_i\beta - \eta_i) - \frac{1}{2}h_\eta \sum_i \eta_i^2 \right] \quad (2.17)$$

2.3   Estimation

We implement a Markov chain Monte Carlo (MCMC) algorithm for the estimation of the parameters and inefficiencies. Specifically, a procedure known as *Metropolis within Gibbs* (see for instance Robert and Casella (2010, p. 230)). The Gibbs sampling algorithm is

used to estimate the parameters for which we have a standard form for the conditional posterior distributions (β, h_μ, and h_η); while the Metropolis-Hasting algorithm is used to estimate the parameters with no standard form; that is, γ, δ, and η_i. For a comprehensive implementation of the Gibbs sampling and the Metropolis-Hastings algorithms see Greenberg (2008, p. 90-102) and for a detailed exposition of the Metropolis-Hastings algorithm see Chib and Greenberg (1995).[11]

2.4 Simulation Exercises

We implement some simulation exercises to validate the performance of the proposed method. We generate synthetic data for three different scenarios–a small, a medium, and a large data set. The small data set simulates information for 50 companies, the medium size data set simulates data for 250 companies, and the large data set simulates data for 1,000 companies. Each data set follows the likelihood structures presented in Section 2.1 and use as population parameters $\beta = \gamma = \delta = h_\mu = h_\eta = 1$. We run the Monte-Carlo simulation for 20,000 iterations. We disregard the first 2,000 iterations to avoid initial values biases and a thin parameter of 10 to mitigate autocorrelation.

Figure 2.1 presents the marginal posterior distributions of the scale parameters (h_μ and h_η) and the parameter accompanying the inefficiency (δ). As expected, the population parameters fall in the region of high posterior probability mass. Figure 2.2, present the estimated inefficiencies plotted against the true values generated according to subsection 2.1.1; e.g., $\eta_i \sim \mathcal{N}^+\left(0, h_\eta^{-1}\right)$. What this plot tells us is that the maximum likelihood estimates are more dispersed than the Bayesian ones, and as so, the Bayesian technique provides better estimates. In a perfect scenario, the scatter of the estimates and the population parameters should form a $45°$ line.

Besides our proposed estimation method, we also estimate the parameters by maximum likelihood following the two-stage approach common in the literature. Both estimations are repeated 100 times for each synthetic data set. We compute different error measures which are standard in the literature when comparing different estimation methods. Specifically, we compute: the mean absolute error (MAE), the mean percentage absolute error (MPAE), the mean root squared error (MRSE), and the mean squared error (MSE). The results for our parameter of interest δ, the parameter accompanying the inefficiency, are presented in Table 2.1.[12]

As expected, the proposed simultaneous estimation method has smaller estimation errors regardless of the error measure (MAE, PAE, RSE, MSE). However, the advantage of the simultaneous estimation fades away as the size of the data set increases. As a drawback of our proposed method, we see in Table 2.1 that increasing the sample size from 250 to 1,000 increases a bit the mean estimation error, though it is still smaller than the errors if we use two stages maximum likelihood. This result might be because with a data set

---

[11]All computations are made in the statistical software R (2017) and are available upon request.

[12]The results for the other parameters are available upon request.

Figure 2.1: **Marginal Posterior Distributions of Scale Parameters and Inefficiency**

(a) h_μ (b) h_η (c) δ
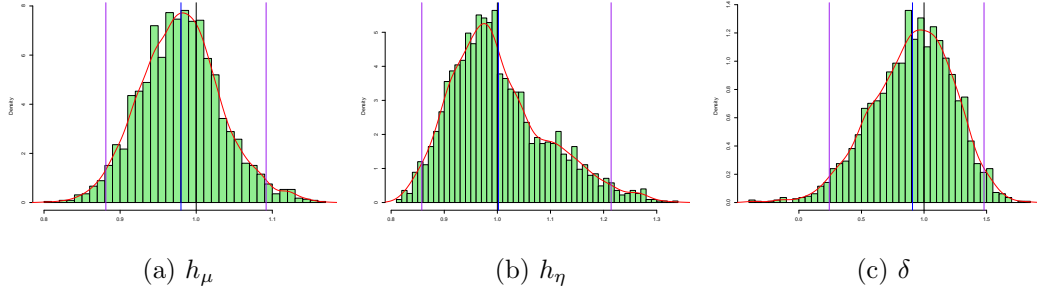
Note: This plot presents the marginal posterior distribution of the scale parameters (h_μ and h_η) and the parameter of the inefficiency (δ). As expected, the population parameters ($h_\mu = h_\eta = \delta = 1$) fall in the region of high posterior probability mass depicted between the two outer vertical lines.

Figure 2.2: **Scatterplot of Bayesian and ML Estimated of Inefficiencies**

(a) Bayesian estimates   (b) Maximum likelihood estimates
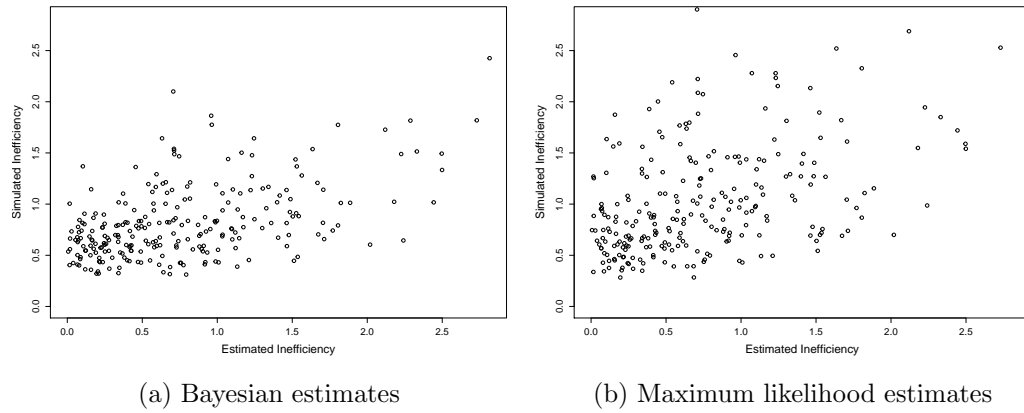
Note: This plot presents the scatterplots of the Bayesian and maximum likelihood estimates vs the simulated inefficiencies.

10

Table 2.1: **Simulation Results**

	MAE	PAE	RSE	MSE
Sample set = 50				
Joint Bayesian	0.54186	0.54186	0.72645	0.52772
Two-stage MLE	51.98920	33.75922	201.19973	40,481.33151
Sample set = 250				
Joint Bayesian	0.20859	0.20859	0.24809	0.06155
Two-stage MLE	17.70118	11.49428	49.07837	2,408.68674
Sample set = 1,000				
Joint Bayesian	0.33877	0.33877	0.34185	0.11686
Two-stage MLE	1.27013	0.82476	6.34006	40.19640

Note: This table shows the Mean Absolute Error (MAE), the Percentage Absolute Error (PAE), the Root Squared Error (RSE), and Mean Squared Error (MSE) for three different scenarios. All population parameters are set to one and the models presented in subsection 2.2.2 are estimated by the proposed Bayesian method and the two-stage approach by maximum likelihood (MLE) as it is common in the literature.

of 1,000 firms we are in fact estimating about 1,005 parameters–1,000 inefficiencies plus the other covariates (β, γ, δ, h_η, h_μ); while by maximum likelihood one just estimates the covariates and computes the inefficiencies as:[13]

$$\mathbb{E}\left[\eta_i | \epsilon_i\right] = \mu_i^* + \frac{\sigma^* \phi\left(\frac{-\mu_i^*}{\sigma^*}\right)}{1 - \Phi\left(\frac{-\mu_i^*}{\sigma^*}\right)} \tag{2.18}$$

where $\epsilon_i = \mu_i + \eta_i$, and

$$\mu_i^* = \frac{\sigma_\eta^2 \epsilon_i}{\sigma_\mu^2 + \sigma_\eta^2} \qquad \text{and} \qquad \sigma^* = \sqrt{\frac{\sigma_\eta^2 \sigma_\mu^2}{\sigma_\mu^2 + \sigma_\eta^2}} \tag{2.19}$$

3 Data

Our data set is composed of the financial information of all U.S. commercial banks filling the quarterly Reports of Condition and Income (Call Reports).[14] All nominal values are in

---

[13]Kumbhakar et al. (2015, p. 67-68) present the derivation of the technical efficiency for a production function. To derive the technical efficiency of the a cost function one just needs to change a couple of things as the authors explain in page 117.

[14]The Call reports, up to 2010, can be downloaded from the Federal Reserve Bank of Chicago from https://www.chicagofed.org/banking/financial-institution-reports/commercial-bank-data-complete-2001-2010.

thousands of U.S. Dollars and are deflated using the 2005 Consumer Price Index published by the Bureau of Labor Statistics.[15] The information for the failed banks comes from the Federal Deposit Insurance Corporation (FDIC).[16]

For our application we use the last available financial statement prior to the bank's failure. The time period is from 2001 to 2010. The reason for this period is because it comprises the financial crisis, which came with a spike in the number of bank failures (see Figure 3.1). Consequently, it serves as a natural experiment for the problem we want to analyze. Figure 3.1 plots the number of bank failures per year.

Figure 3.1: **Number of Failed Commercial Banks in U.S.**
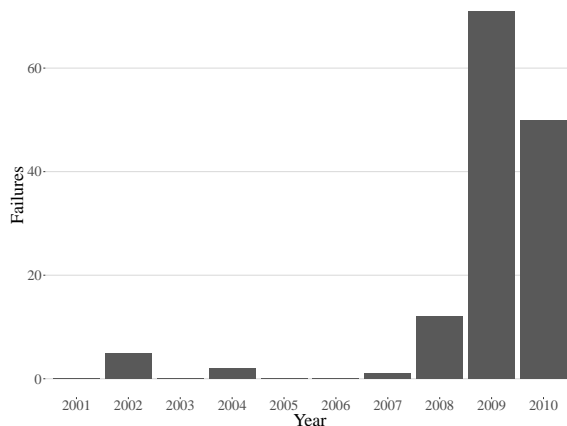
For the stochastic frontier model and as it is common in the literature, we delete the data with either missing or negative values in cost, equity, and total assets, as well as in the input prices and output quantities. For the variables construction, we follow Koetter, Kolari, and Spierdijk (2012) (see Table 3.1).

For the estimation of the stochastic frontier model, we follow Kumbhakar et al. (2015, p. 103-104) and assume a translog functional form for the cost function (Equation 3.1).

---

[15]The data can be downloaded from https://data.bls.gov/pdq/SurveyOutputServlet

[16]The data can be downloaded from: https://www5.fdic.gov/hsob/SelectRpt.asp?EntryTyp=30

Table 3.1: **Stochastic Frontier Model Variables**

Name and Acronym		Description
Total cost	$Cost$	Sum of interest expenses on deposits (riad4170), interest expenses on fed funds (riad4180), loan-loss provisions (riad4230), expenditures on fixed assets (riad4217), and salaries (riad4135)
Cost of fixed assets	w_1	Expenditures on fixed assets (riad4217) divided by the premises and fixed assets (rcfd2145)
Cost of labor	w_2	Salaries (riad4135) divided by full-time equivalent employees (riad4150)
Cost of borrowed funds	w_3	Interest expense on deposits (riad4170) plus interest expense on fed funds (riad4180) divided by the sum of total deposits (rcfd2200) and fed funds purchased (rcfd2800)
Total securitues	y_1	Securities held to maturity (rcfd1754) and securities held for sale (rcfd1773)
Total loans	y_2	Total loans and leases (rcfd1400)

Note: This table shows the stochastic frontier model's variables and their construction following Koetter et al. (2012). Source: This table is an adaptation of Table 1 in Koetter et al. (2012, p. 467).

$$
\begin{aligned}
\operatorname{Ln}\left(\frac{C_i^a}{w_{3i}}\right) = {} & \beta_0 + \beta_1 \operatorname{Ln}(y_{1i}) + \beta_2 \operatorname{Ln}(y_{2i}) + \beta_3 \operatorname{Ln}(y_{1i}) \operatorname{Ln}(y_{2i}) + \beta_4 \operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right) + \beta_5 \operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right) + \\
& \beta_6 \operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right) \operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right) + \frac{1}{2}\beta_7 \left[\operatorname{Ln}(y_{1i})\right]^2 + \frac{1}{2}\beta_8 \left[\operatorname{Ln}(y_{2i})\right]^2 + \frac{1}{2}\beta_9 \left[\operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right)\right]^2 + \\
& \frac{1}{2}\beta_{10} \left[\operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right)\right]^2 + \beta_{11} \operatorname{Ln}(y_{1i}) \operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right) + \beta_{12} \operatorname{Ln}(y_{1i}) \operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right) + \\
& \beta_{13} \operatorname{Ln}(y_{2i}) \operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right) + \beta_{14} \operatorname{Ln}(y_{2i}) \operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right) + \mu_i + \eta_i
\end{aligned}
$$

$$(3.1)$$

For the proportional hazards model, we follow Alvarez Franco and Restrepo Tobon (2016) and construct the variables that best proxy the factors used by the FDIC's CAMELs rating. Besides the advantage of comparability with previous literature, different articles show that these variables capture the banks's overall condition, as well as their probability of failure (see for instance A. Berger, Imbierowicz, and Rauch (2016); Cole and White (2012); Wheelock and Wilson (2000)). The variables construction are reported in Table 3.2.

Table 3.3 presents summary statistics of the main financial variables of the commercial banks. From this table, we can see that there is a large heterogeneity of the banks in the sample. There are small banks with assets of sixteen millions while others have assets of more than 4.5 billions. Likewise, there are banks with about ten millions in deposits and loans while others have between three and four billions. Analyzing the profits, we find that more than 2.5% of the sample have negative profits. In the last percentiles, we see a remarkable spike in profits. While up to 75% percent of firms have profits of less than

Table 3.2: **Proportional Hazards Model Variables**

Capital adequacy		
C_1	Equity / Assets	Equity (rcfd3210), Assets (rcfd2170)
C_2	Tier 1 capital	Tier 1 capital (rcfd8274)
C_3	Tier 2 capital	Tier 2 capital (rcfd5311)
Assets quality		
A_1	Loans / Assets	Loans (rcfd2122), Assets (rcfd2170)
A_2	Real estate loans / Loans	Real estate loans (rcfd1410), Loans (rcfd2122)
A_3	Commercial and industrial loans / Loans	Commercial and industrial loans (sum of rcfd1763 plus rcfd1764), Loans (rcfd2122)
A_4	Loan loss provision /Loans	Loan loss provision (riad4230), Loans (rcfd2122)
A_5	Non-performing loans / Loans	Non-performing loans (rcfd5613), Loans (rcfd2122)
A_6	Loan loss provision / Assets	Loan loss provision (riad4230), Assets (rcfd2170)
A_7	Non-performing loans / Assets	Non-performing loans (rcfd5613), Assets (rcfd2170)
A_8	Off-balance sheet activities / Assets	Off-balance sheet activities (ratio of riad4079 and riad4000), Assets (rcfd2170)
Earnings quality		
E_1	Profits / Equity	Profits (riad4000 minus Cost), Equity (rcfd3210)
E_2	Profits / Assets	Profits (riad4000 minus Cost), Assets (rcfd2170)
Liquidity		
L_1	Loans / Deposits	Loans (rcfd2122), Deposits (rcfd2200)

Note: This table shows the proportional hazards model's variables and their construction following Alvarez Franco and Restrepo Tobon (2016).

6.5 million, the top 2.5% have profits of more than 33.5 millions and a maximum of 70.5 millions.

Table 3.3: **Summary Statistics for the Financial Variables**

				Percentiles					
Vars	**Min**	**Mean**	**sd**	**2.5^{th}**	**25th**	**50th**	**75th**	**97.5^{th}**	**Max**
Assets	16,415.74	311,504.69	405,178.52	30,875.13	91,095.77	173,850.90	353,314.28	1,539,159.46	4,610,134.00
Equity	990.45	31,154.77	45,915.59	2,881.35	8,823.18	16,804.51	34,202.91	154,044.29	981,253.00
Deposits	14,764.25	255,689.98	322,471.39	26,459.19	76,822.61	145,876.40	293,929.26	1,234,870.86	3,902,502.94
Loans	10,402.51	205,430.56	270,945.34	16,603.71	55,921.47	112,486.03	234,631.15	995,971.50	2,878,926.00
RE.loans	0.00	156,589.96	214,092.03	8,057.01	37,207.81	83,153.41	181,719.44	768,449.24	2,571,644.13
Prov.loans	8.81	2,201.48	4,180.26	23.14	201.87	641.64	2,041.78	16,318.30	31,879.58
Cost	395.00	11,707.93	14,865.34	1,096.85	3,301.07	6,490.28	13,501.82	55,892.49	154,972.35
Off.bal	-0.10	0.11	0.07	0.00	0.06	0.10	0.15	0.29	0.42
Profits	-26,973.59	5,545.05	9,140.78	-3,932.31	1,162.32	2,830.34	6,463.01	33,579.74	70,573.76

Note: This table shows the summary statistics for the financial variables of the commercial banks in the data set. All nominal variables are in thousands of U.S. Dollars and are deflated using the 2005 Consumer Price Index published by the Bureau of Labor Statistics.

Table 3.4 presents the summary statistics for the variables used in the stochastic frontier model. The total cost, which is our dependent variable, is relatively dispersed. Some banks have a total cost of just 395 thousands while others have a total cost of more than 150 millions. The most representative output is loans (y_2) with a minimum of 10.5 millions while securities (y_1) is just 350 thousand in its minimum.

Table 3.5 presents summary statistics for the variables for the hazards model. In the capital adequacy variable (C_1) we see that some banks have less than 1% of Equity in terms of Assets while other have almost 60%. The assets quality variables (A_1, A_2, A_4, A_6) show that some banks have a proportion of loans of 95% of their assets (A_1) and that some banks have all of their loans as real estate loans (A_2). However, the provisions for loan losses in terms of their assets (A_6) is less than 10%. Also, call the attention the liquidity quality variable (L_1) from which we can see that there are banks with loans of two and a

Table 3.4: **Summary Statistics for the Stochastic Frontier Model**

Vars	Min	Mean	sd	Percentiles 2.5$^{\text{th}}$	25$^{\text{th}}$	50$^{\text{th}}$	75$^{\text{th}}$	97.5$^{\text{th}}$	Max
$Cost$	395.00	11,707.93	14,865.34	1,096.85	3,301.07	6,490.28	13,501.82	55,892.49	154,972.35
y_1	358.06	64,003.40	105,609.79	1,797.36	12,599.10	29,327.43	67,300.66	356,299.83	1,595,729.36
y_2	10,460.90	205,533.59	271,054.20	16,650.49	55,950.94	112,566.48	234,711.75	995,971.50	2,878,926.00
w_1	0.05	0.35	0.40	0.08	0.15	0.22	0.37	1.57	3.33
w_2	5.96	65.92	19.95	36.01	53.55	62.56	74.51	115.53	260.10
w_3	0.00	0.02	0.01	0.01	0.01	0.01	0.02	0.04	0.08

Note: This table shows the summary statistics for the variables to be considered in the estimation of the stochastic frontier model. All variables are in thousands of U.S. Dollars and are deflated using the 2005 Consumer Price Index published by the Bureau of Labor Statistics. The variables are: Total cost (the dependent variable), Total securities (y_1), Total loans (y_2), Cost of fixed assets (w_1), Cost of labor (w_2), and Cost of borrowed funds (w_3). Refer to Table 3.1 and Koetter et al. (2012) for further details about the variables construction.

half times their deposits.

Table 3.5: **Summary Statistics for the Hazards Model**

Vars	Min	Mean	sd	Percentiles 2.5$^{\text{th}}$	25$^{\text{th}}$	50$^{\text{th}}$	75$^{\text{th}}$	97.5$^{\text{th}}$	Max
$Camel_{C1}$	0.0092	0.1016	0.0322	0.0519	0.0838	0.0957	0.1133	0.1791	0.5848
$Camel_{A1}$	0.0761	0.6506	0.1348	0.3406	0.5672	0.6685	0.7483	0.8682	0.9527
$Camel_{A2}$	0.0000	0.7248	0.1669	0.3280	0.6322	0.7539	0.8478	0.9694	1.0000
$Camel_{A4}$	0.0001	0.0101	0.0130	0.0005	0.0026	0.0056	0.0119	0.0487	0.1303
$Camel_{A6}$	0.0000	0.0068	0.0090	0.0003	0.0016	0.0035	0.0078	0.0342	0.0840
$Camel_{E1}$	-1.2586	0.1644	0.1712	-0.2816	0.1215	0.1904	0.2518	0.3932	0.5010
$Camel_{L1}$	0.0921	0.7802	0.1788	0.4022	0.6715	0.7926	0.8960	1.1000	2.5786

Note: This table shows the summary statistics for the variables to consider in the estimation of the proportional hazards model. The variables are: Capital adequacy ($C_1 = Equity/Assets$); Assets quality ($A_1 = Loans/Assets$, $A_2 = RE.loans/Loans$, $A_4 = Prov.loans/Loans$, $A_6 = Prov.loans/Assets$); Earnings quality ($E_1 = Profits/Equity$), and Liquidity quality ($L_1 = Loans/Deposits$). Refer to Table 3.2 and Alvarez Franco and Restrepo Tobon (2016) for further details about the variables construction.

To further understand our data set, we compute the Welch's test to check if there is any difference in mean between the failing firms and the non-failing ones. Explicitly, we compare the mean of our variables of interest like output levels (y_1 and y_2), cost (w_1, w_2, and w_3), CAMELs, and the efficiency and inefficiency. For this exercise, we first estimate the inefficiencies for each firm following equation 2.17.

According to the results (see Table 3.6), the failing firms have more loans (y_2), spend more in salaries (w_2), have more assets, and are more inefficient than the non-failing firms. Likewise, failing firms have fewer securities (y_1) and less ratio equity to assets (C_1) than non-failing firms. The fact that the failing firms are more inefficient give us some indication that the level of inefficiency plays a role in determining the time to failure.

Table 3.6: **Difference in Mean: Welch's Test**

Variable	Failing firms	Non-failing firms	t-statistic	
y_1	36,568.32	64,558.96	3.61	*
y_2	260,383.03	204,422.90	8.31	*
w_1	0.41	0.35	12.53	*
w_2	77.71	65.68	11.92	*
w_3	0.03	0.02	2.80	*
C_1	0.06	0.10	3.74	*
A_1	0.74	0.65	15.91	*
A_2	0.80	0.72	19.23	*
A_4	0.04	0.01	1.73	
A_6	0.03	0.01	1.64	
E_1	-0.31	0.17	-0.28	
L_1	0.85	0.78	21.85	*
Assets	342,301.64	310,881.06	20.79	*
Inefficiency	0.49	0.35	5.85	*
Efficiency	0.63	0.72	15.24	*

Note: This table shows the Welch's test for difference in mean between the failing and non-failing firms. The asterisk (*) represents statistical significance at 5%.

4 Empirical Results

Table 4.1 presents the results of the parameter estimates. The first column in the left is the mean (usually regarded as the point estimate) of the Markov chains for the simultaneous Bayesian estimation method. The second, third, and fourth columns are the 2.5, 50, and 97.5 percentiles of the chains. This gives us the 95% credible intervals for the Bayesian estimates. The last two columns report the estimates by maximum likelihood and the associated t-statistic at 5% significance level.

We check the robustness of our results by using different sets of covariates. For instance, we use the average of the last five available financial statements, instead of the last available prior to the bank failure. We also repeat all estimations without deflating data set with the consumer price index. All results are similar to the ones presented here (see Appendix B Table B.1).

In the estimation of the stochastic frontier, both methods obtain the same sign (either positive or negative) for all parameters $(\beta_i, \sigma_\mu, \sigma_\eta)$. In the estimation of the proportional hazards model, we also obtain congruent signs in most of the parameters.

Our parameter of interest δ is positive and statistically significant in the simultaneous estimation. A reason for which this parameter might not result statistically significant in the maximum likelihood estimation is that the two-stage approach produces inefficient estimators and consequently larger standard errors. This leads to higher probabilities to fall in the no-rejection zone of the null-hypothesis of no significance of such parameter.

To better understand the effect of the inefficiency over the probability of failure, recall that we are modeling the time to failure assuming an exponential model. That is, $y \sim \text{Exp}(\lambda)$ and $\lambda = \exp(z'\gamma + \delta\eta)$, where λ is the failure rate. By properties of the exponential distribution, the mean time to failure equals the inverse of the failure rate. In other words, $\mathbb{E}[y] = \lambda^{-1}$. Now, a positive δ means that increasing the inefficiency increases the failure rate and consequently reduces the mean time to failure. i.e. higher inefficiency implies a higher probability of failure. This result is consistent with economic intuition and with previous research

If we express the hazards rate as $\lambda = \exp(z'\gamma)\exp(\delta\eta)$, recall that Efficiency $= \exp(-\eta)$, for simplicity assume that $\exp(z'\gamma) = 1$, we can relate the mean time to failure and the efficiency (see Figure 4.1). What this exercise tell us is that having an efficiency rate of 0% (complete inefficiency) implies a proportional mean time to failure of about 20%; while being fully efficient implies a proportional mean time to failure of 100%. However, the relation is not linear. Increasing the efficiency from 2% to 3% (one percentage point) increases the firms' proportional mean time to failure to about 30% (four percentage points), while if the improvement in efficiency is from 90% to 91%, the increase in the proportional mean time to failure is of just 0.3 percentage points.

The analysis with the other covariates of the hazards model is straightforward. Having a higher ratio of Loans/Assets (A_1), Real Estate Loans/Loans (A_2), and Loans/Deposits (L_1) decreases the mean time to failure. On the other hand, having a higher ratio of

17

Figure 4.1: **Efficiency vs Mean Time to Failure**
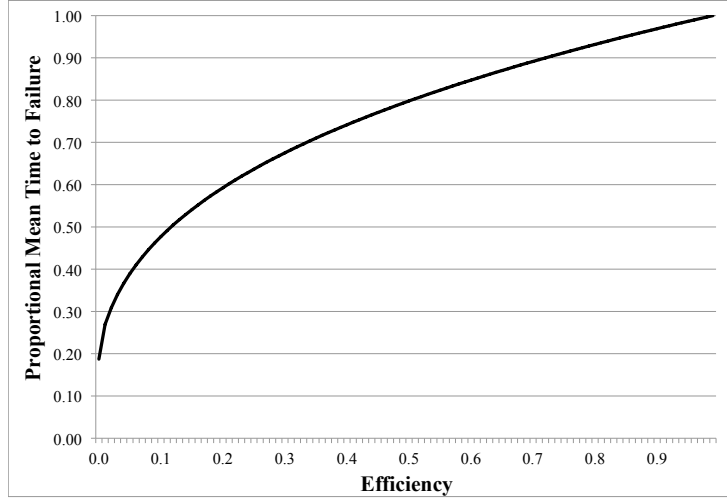
Equity / Assets (C_1), Loan loss provision/Loans (A_4), Profits/Equity (E_1), and Assets (Ln *Assets*) increases the mean time to failure.

Most of our results are consistent with the ones presented in Almanidis and Sickles (2016); Alvarez Franco and Restrepo Tobon (2016); though there are some that are not, these are not statistically significant. We compare our results with these articles because they use similar data and time period.

4.1 Cost Elasticities

One of the advantages of the Bayesian estimation is that we can perform easily statistical inference of the cost elasticity in terms of the total securities (y_1) as in equation 4.1. The computation of the cost elasticity in terms of total loans (y_2), in terms of the relative cost of fixed assets (w_1/w_3), and in terms of the relative cost of labor (w_2/w_3) is similar to the one presented here.

$$
\begin{aligned}
\varepsilon_{y_{1i}} &= \frac{\partial \operatorname{Ln}\left(\frac{C_i^a}{w_{3i}}\right)}{\partial \operatorname{Ln}(y_{1i})} \\
&= \beta_1 + \beta_3 \operatorname{Ln}(y_{2i}) + \beta_7 \operatorname{Ln}(y_{1i}) + \beta_{11} \operatorname{Ln}\left(\frac{w_{1i}}{w_{3i}}\right) + \beta_{12} \operatorname{Ln}\left(\frac{w_{2i}}{w_{3i}}\right)
\end{aligned}
\tag{4.1}
$$

Since we have the posterior chains of each parameter, we can obtain an empirical distribution for each of the cost elasticities for each firm. Figure 4.2 presents the histograms of the cost elasticities in terms of the outputs and relative cost of inputs. According to the
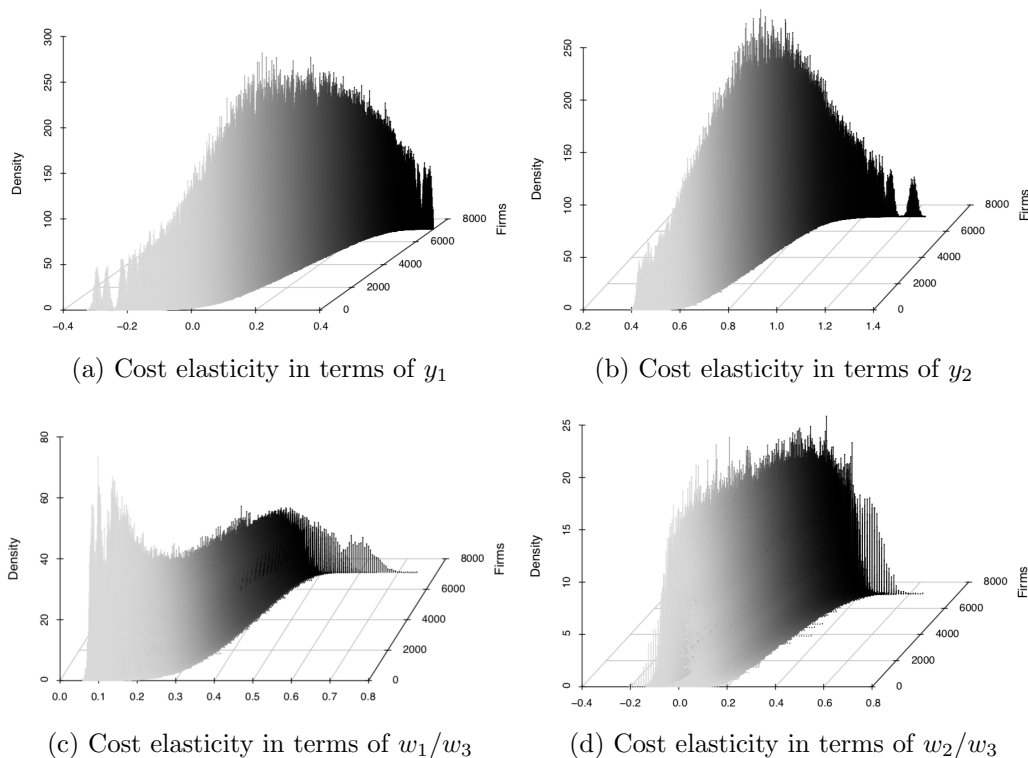
Table 4.1: **Estimation results**

	Bayesian estimation				ML estimation		
	Mean	**Perc. 2.5$^{\text{th}}$**	**Perc. 50$^{\text{th}}$**	**Perc. 97.5$^{\text{th}}$**	**Estimates**	**t-test**	
$Cons$	2.3643	1.7919	2.3632	2.9109	3.1972	3.4869	*
Ln(y_1)	0.1870	0.1399	0.1866	0.2365	0.1635	1.6372	
Ln(y_2)	0.6094	0.5418	0.6096	0.6792	0.5454	3.7423	*
Ln(y_1)Ln(y_2)	-0.0725	-0.0759	-0.0725	-0.0690	-0.0767	-9.9572	*
Ln(w_1/w_3)	-0.2419	-0.3145	-0.2416	-0.1731	-0.2191	-2.3730	*
Lnw_2/w_3	-0.2062	-0.3133	-0.2083	-0.0946	-0.2413	-1.3852	
Ln(w_1/w_3)Ln(w_2/w_3)	0.0498	0.0410	0.0496	0.0587	0.0459	3.7739	*
Ln(y_1)^2	0.0681	0.0649	0.0681	0.0716	0.0675	9.6133	*
Ln(y_2)^2	0.0902	0.0840	0.0902	0.0962	0.1003	10.1324	*
Ln(w_1/w_3)^2	-0.0655	-0.0727	-0.0655	-0.0581	-0.0674	-6.7332	*
Ln(w_2/w_3)^2	0.0824	0.0677	0.0824	0.0974	0.0817	4.1543	*
Lny_1Ln(w_1/w_3)	0.0088	0.0051	0.0088	0.0128	0.0175	3.6793	*
Lny_1Ln(w_2/w_3)	0.0082	0.0028	0.0082	0.0134	0.0132	1.3448	
Lny_2Ln(w_1/w_3)	-0.0032	-0.0081	-0.0031	0.0019	-0.0086	-1.1813	
Lny_2Ln(w_2/w_3)	-0.0205	-0.0275	-0.0205	-0.0133	-0.0221	-1.3253	
σ_μ	0.1132	0.1085	0.1134	0.1182	0.1910	67.6902	*
σ_η	0.4241	0.4154	0.4242	0.4323	0.4748	42.0861	*
Camel C_1	-0.8951	-1.3528	-0.9481	-0.3024	-1.7093	-0.7440	
Camel A_1	0.3335	-0.1580	0.3607	0.8531	-2.9353	-1.8433	
Camel A_2	0.8491	0.5460	0.8222	1.3124	5.1231	6.9988	*
Camel A_4	-0.0243	-0.4028	-0.0056	0.3351	-1.3209	-0.0767	
Camel A_6	0.2379	-0.3378	0.3256	0.6807	3.5461	0.1415	
Camel E_1	-3.1125	-3.5663	-3.3133	-1.4301	-3.0999	-10.1207	*
Camel L_1	0.8802	0.1849	0.9969	1.3206	1.5193	1.2945	
Ln $Assets$	-0.6514	-0.7136	-0.6567	-0.5690	-0.7880	-11.4693	*
Ineff η	0.3351	0.0176	0.3249	0.6509	0.3079	0.6262	

Note: This table shows the results of the estimation by the proposed simultaneous Bayesian method and the two-stage approach by maximum likelihood. The estimates with asterisk are statistically significant at 5% confidence level. Our parameter of interest (δ) is significant and it has the expected positive sign. It means that being more inefficient increases the failure function rate and consequently decreases the mean time to failure. That is, it increases the probability of failure.

results, most of the elasticities are between zero and one. This means that increasing in 1% any of the outputs (y_1, y_2) or the relative cost (w_1/w_3, w_2/w_3) increases the total cost in less than 1%. In the case where the cost elasticity is higher than one, we have that 1% increase in either outputs or relative cost, increases in more than 1% the total cost of the specific firm.

Figure 4.2: **Histograms and Empirical Density of Cost Elasticities**

(a) Cost elasticity in terms of y_1

(b) Cost elasticity in terms of y_2

(c) Cost elasticity in terms of w_1/w_3

(d) Cost elasticity in terms of w_2/w_3

Note: This plot presents the histograms of the cost elasticity in terms of total securities (y_1), total loans (y_2), the relative cost of fixed assets $\left(\frac{w_1}{w_3}\right)$, and the relative cost of labor $\left(\frac{w_2}{w_3}\right)$.

Figure 4.2 also allows us to appreciate the difference in cost elasticities between firms. For instance, while some firms have cost elasticity close to zero, other ones have their cost elasticity close or even bigger than one. Besides, we can appreciate the uncertainty associated with each banks elasticity.
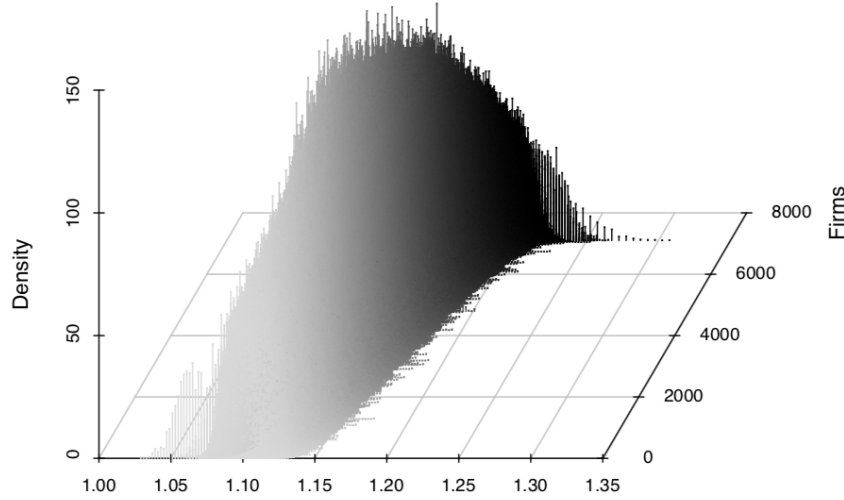
4.2 Returns to Scale

Similarly to the cost elasticity, the returns to scale measures the percentage increase in total cost given a proportional increase in *all* outputs. Algebraically,

$$RTS = \left[\frac{\partial \frac{C_i^a}{w_3}}{\partial y_1} + \frac{\partial \frac{C_i^a}{w_3}}{\partial y_2}\right]^{-1} \tag{4.2}$$

Figure 4.3 presents the plot of the histograms for the returns to scale per firm. In addition to the differences between banks that we mention above; we can also appreciate that any of the banks in our sample have increasing returns to scale ($RTS > 1$). This means that increasing all outputs in one percent increases total cost in less than one percent. Figure 4.3 also allows seeing the banks heterogeneity. There is a bank with returns to scale of about 1.05 while another one with returns to scale of around 1.25. This means that if both banks increase in one percent both outputs, the first bank will have an increase in its cost of 0.95% (1/1.05) while the second one will have an increase in the cost of about 0.8%.

Figure 4.3: **Returns to Scale**

Note: This plot presents the histograms of the returns to scale for each bank.

4.3 Distribution of Efficiencies per Firms

In the Bayesian estimation, each bank's inefficiency is treated as an unknown parameter. Consequently, we obtain conditional posterior distributions for each inefficiency. This is not possible using maximum likelihood estimation because the estimation of the inefficiency is obtain using equation 2.18 which gives a point estimate per firm. Instead, we have a conditional posterior distribution per firm. Figure 4.4 presents examples of estimated

efficiencies. We can see that while Bank "2129" has an efficiency ranging from 60% to 70%; Bank "6165" has an efficiency between 80% to around 100%.

Figure 4.4: **Empirical Densities of Efficiencies**
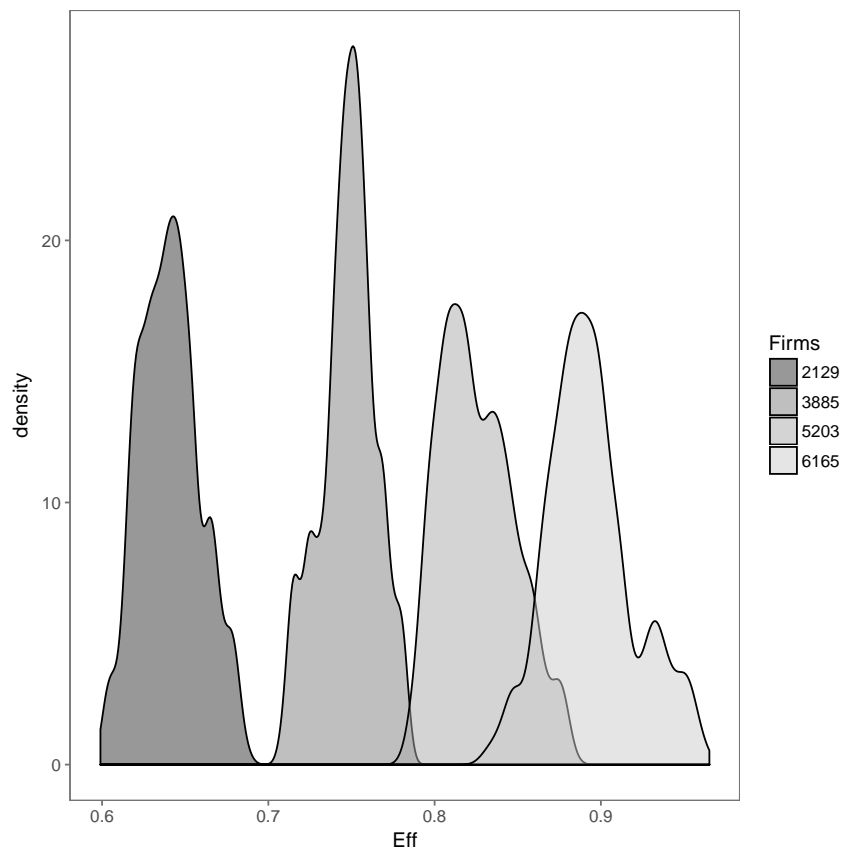
Note: This plot presents the empirical densities of the efficiency for four different firms.

5   Concluding Remarks

Several authors attempt to estimate the effect of inefficiency on the probability of failure. Since the inefficiency is not observable, the usual approach consists of estimating the inefficiency in a first step to then include the results as a covariate in the estimation of the probability of failure. This approach, however, yields inefficient, biased, and inconsistent estimators because the second step ignores the measurement error of the first step.

We propose to solve this problem by estimating, simultaneously, a stochastic frontier

model and a hazards model. Our proposed method consists of estimating the parameters by Gibbs-Sampling within Metropolis-Hastings using the conditional posterior distributions of the parameters. The choice of this estimation method is because we can use the conditional posterior distributions of the parameters to perform statistical inference of functions of the parameters like elasticities, returns to scale. Besides, we can treat each bank inefficiency as an unknown parameter and as such obtain its respective conditional posterior distribution.

Through simulation exercises, we show that, effectively, the estimation error of the two-stage approach, via maximum likelihood, is higher than the one observed if we use the simultaneous Bayesian estimation.

Using commercial banks data from 2001 to 2010, we find that, as expected, more inefficient banks have a lower mean time to failure. In addition, having higher ratios of loans to assets and real estate loans to total loans increases the probability of failure (fail sooner).

Finally, we compute the cost elasticity in terms of the total securities (y_1), total loans (y_2), the relative price of fixed assets (w_1/w_3), and the relative price of labor (w_2/w_3). We find that most banks have cost elasticities between zero and one. This means that increasing in 1% one of the outputs or the relative prices, increases the total cost in less than one percent. Likewise, computing the returns to scale we find that banks present increasing returns to scale. That is, increasing all outputs in 1% increase the total cost in less than one percent.

In the computation of the cost efficiencies and the returns to scale, we can appreciate one of the advantages of the Bayesian estimation. Since we have the posterior chains of the estimates, we can compute the cost efficiencies and the returns to scale per firm acknowledging for the uncertainty associated with the parameter estimates.

In terms of economic policy, the results indicate that bigger banks, measured by their total assets, have a smaller probability of failure and also have increasing returns to scale. Hence, to promote market efficiency, policy maker and regulators should not heavily restrict mergers and acquisitions between commercial banks. On the other hand, as efficiency is a statistically significant determinant of bank failures, regulators should consider within their inspections the cost efficiency level of the bank under analysis in order to reduce its probability of failure.

References

Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, *6*(1), 21–37. doi: 10.1016/0304-4076(77)90052-5

Almanidis, P., & Sickles, R. C. (2016). Banking Crises, Early Warning Models, and Efficiency. In J. Aparicio, C. A. K. Lovell, & J. T. Pastor (Eds.), *Advances in efficiency and productivity* (1st ed., pp. 331 – 364). Springer International Publishing. doi: 2451/10.1007/978-3-319-48461-7_14

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, *23*(4), 589–609. doi: 10.2307/2978933

Alvarez Franco, P. B., & Restrepo Tobon, D. A. (2016). Managerial Efficiency and Failure of U.S. Commercial Banks During the 2007:2009 Financial Crisis: Was this time different? *Ecos de Economía*, *20*(43), 4–22. doi: 10.17230/ecos.2016.43.1

Assaf, A., Berger, A. N., Roman, R. A., & Tsionas, E. G. (2017). Does Efficiency Help Banks Survive and Thrive during Financial Crises? *SSRN Electronic Journal*. doi: 10.2139/ssrn.2970251

Barr, R. S., Seiford, L. M., & Siems, T. F. (1993). An Envelopment-Analysis Approach to Measuring the Managerial Efficiency of Banks. *Annals of Operations Research*, *45*(1), 1–19. doi: 10.1007/BF02282039

Barr, R. S., & Siems, T. F. (1997). Bank Failure Prediction Using DEA to Measure Management Quality. In *Interfaces in computer science and operations research* (pp. 341 – 365). Boston, Massachusetts: Springer U.S. doi: 10.1007/978-1-4615-4102-8_15

Battese, G. E., & Corra, G. S. (1977). Estimation of a Production Frontier Model: with Application to the Pastoral Zone of Eastern Australia. *Australian Journal of Agricultural Economics*, *21*(3), 169 – 179. Retrieved from http://ageconsearch.umn.edu/handle/22266

Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, *4*, 71–111. doi: 10.2307/2490171

Berger, A., Imbierowicz, B., & Rauch, C. (2016). The Roles of Corporate Governance in Bank Failures During the Recent Financial Crisis. *Journal of Money, Credit and Banking*, *48*(4), 729–770. doi: 10.1111/jmcb.12316

Berger, A. N., & Humphrey, D. B. (1992). Measurement and Efficiency Issues in Commercial Banking. In Zvi Griliches (Ed.), *Output measurement in the service sectors* (pp. 245 – 300). University of Chicago Press. Retrieved from http://www.nber.org/chapters/c7237.pdf

Brewer III, E., Genay, H., Hunter, W. C., & Kaufman, G. G. (2003). The Value of Banking Relationships During a Financial Crisis: Evidence from Failures of Japanese Banks. *Journal of the Japanese and International Economies*, *17*(3), 233–262. doi: 10.1016/S0889-1583(03)00024-8

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics*. Cambridge University Press.

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, *49*(4), 327–335. doi: 10.2307/2684568

Cole, R. A., & Gunther, J. W. (1998). Predicting Bank Failures: A Comparison of On-and Off-Site Monitoring Systems. *Journal of Financial Services Research*, *13*(2), 103 – 117. doi: 10.1023/A:1007954718966

Cole, R. A., & White, L. J. (2012). Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures This Time Around. *Journal of Financial Services Research*, *42*(1), 5– 29. doi: 10.1007/s10693-012-0135-1

Cole, R. A., & Wu, Q. (2009). Is Hazard or Probit Model more Accurate in Predicting Financial Distress? Evidence from U.S. Bank Failures. *MPRA - Munich Personal RePec Archive*(29182). Retrieved from https://mpra.ub.uni-muenchen.de/29182/

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–220. Retrieved from http://www.jstor.org/stable/2985181

Dimara, E., Skuras, D., Tsekouras, K., & Tzelepis, D. (2008). Productive efficiency and firm exit in the food sector. *Food Policy*, *33*(2), 185–196. doi: 10.1016/j.foodpol .2007.08.003

Greenberg, E. (2008). *Introduction to Bayesian Econometrics*. Cambridge University Press.

Halling, M., & Hayden, E. (2006). Bank Failure Prediction: A Two-Step Survival Time Approach. *SSRN Electronic Journal*. doi: 10.2139/ssrn.904255

Koetter, M., Kolari, J. W., & Spierdijk, L. (2012, may). Enjoying the Quiet Life under Deregulation? Evidence from Adjusted Lerner Indices for U.S. Banks. *Review of Economics and Statistics*, *94*(2), 462–480. doi: 10.1162/REST_a_00155

Koop, G. (2003). *Bayesian Econometrics* (Vol. 1). West Sussex, England: John Wiley and Sons Ltd. doi: 10.1017/CBO9781107415324.004

Kumar, P. R., & Ravi, V. (2007). Bankruptcy Prediction in Banks and Firms Via Statistical and Intelligent Techniques – A Review. *European Journal of Operational Research*, *180*(1), 1 – 28. doi: 10.1016/j.ejor.2006.08.043

Kumbhakar, S. C., & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge, U.K.: Cambridge University Press.

Kumbhakar, S. C., Wang, H.-J., & Horncastle, A. (2015). *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. New York, U.S.A.: Cambridge University Press.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, *18*(2), 435–444. doi: 10.2307/2525757

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131. doi: 10.2307/2490395

R, C. T. (2017). *R: A Language and Environment for Statistical Computing.* Vienna, Austria. Retrieved from https://www.r-project.org

Ramirez Hassan, A., & Montoya Blandon, S. (2017). Welfare Gains of the Poor: An Endogenous Bayesian Approach with Spatial Random Effects. *Econometric Reviews*, *4938*, 1–18. Retrieved from http://dx.doi.org/10.1080/07474938.2016.1261062 doi: 10.1080/07474938.2016.1261062

Robert, C. P., & Casella, G. (2010). *Introducing Monte Carlo Methods with R.*

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, *74*(1), 101–124. doi: 10.1086/209665

Tanner, M., & Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, *82*(398), 528–540. doi: 10.2307/2289463

Tsionas, E. G., & Papadogonas, T. A. (2006). Firm Exit and Technical Inefficiency. *Empirical Economics*, *31*, 535 – 548. doi: 10.1007/s00181-005-0045-2

Whalen, G. (1991). A Proportional Hazards Model of Bank Failure: An Examination of Its Usefulness as an Early Warning Tool. *Federal Reserve Bank of Cleveland, Economic Review*, *27*(1), 21–31.

Wheelock, D. C., & Wilson, P. W. (1995). Explaining Bank Failures: Deposit Insurance, Regulation, and Efficiency. *The Review of Economics and Statistics*, *77*(4), 689–700. Retrieved from http://www.jstor.org/stable/2109816

Wheelock, D. C., & Wilson, P. W. (2000). Why do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions. *The Review of Economics and Statistics*, *82*(1), 127–138. Retrieved from http://www.jstor.org/stable/2646678

A   Derivation of Conditional Posterior Distributions

As mentioned in the body of the article, $\eta_i \sim \mathcal{N}^+ \left(0, \, h_\eta^{-1}\right)$. Hence,

$$
\begin{aligned}
\pi\left(\eta|\theta\right) &= \prod_i \frac{2^{\frac{1}{2}} h_\eta^{\frac{1}{2}}}{\pi^{\frac{1}{2}}} \exp\left[-\frac{1}{2} h_\eta \eta_i^2\right] \\
&\propto h_\eta^{\frac{N}{2}} \exp\left[-\frac{1}{2} h_\eta \sum_i \eta_i^2\right]
\end{aligned}
\tag{A.1}
$$

Following the literature, we assume Normal priors for β, γ, and δ. That is,

$$
\beta \sim \mathcal{N}_k\left(B_0, \, \Sigma_0\right) \qquad \gamma \sim \mathcal{N}_j\left(\Gamma_0, \, \Omega_0\right) \qquad \delta \sim \mathcal{N}\left(\Delta_0, \, h_\delta^{-1}\right) \tag{A.2}
$$

and assume Gamma distributions for the error precision parameters h_μ and h_η, e.g.

$$
h_\mu \sim \mathcal{G}\left(\tfrac{s_\mu}{2}, \, \tfrac{v_\mu}{2}\right) \qquad\qquad h_\eta \sim \mathcal{G}\left(\tfrac{s_\eta}{2}, \, \tfrac{v_\eta}{2}\right) \tag{A.3}
$$

A.1   Conditional Posterior Distribution for β

$$
\pi\left(\beta|y, \, c\right) = \mathcal{L}\left(y|\theta, \, \eta\right) \mathcal{L}\left(c|\theta, \, \eta\right) \pi\left(\eta|\theta\right) \pi\left(\theta_{-\beta}\right) \pi\left(\beta\right)
\tag{A.4}
$$

$$
\begin{aligned}
\propto \exp&\left[\sum_i \nu_i\left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right] \\
&h_\mu^{\frac{N}{2}} \exp\left[-\frac{1}{2} h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right)\right] \\
&\pi\left(\eta|\theta\right) \pi\left(\theta_{-\beta}\right) \frac{1}{(2\pi)^{\frac{k}{2}} |\Sigma_0|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\left(\beta - B_0\right)'\Sigma_0^{-1}\left(\beta - B_0\right)\right]
\end{aligned}
\tag{A.5}
$$

$$
\begin{aligned}
\propto \exp&\left[-\frac{1}{2} h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right)\right] \\
&\exp\left[-\frac{1}{2}\left(\beta - B_0\right)'\Sigma_0^{-1}\left(\beta - B_0\right)\right]
\end{aligned}
\tag{A.6}
$$

$$= \exp\left\{-\frac{1}{2}\left[h_\mu \sum_i (c_i + \eta_i)'(c_i + \eta_i) - h_\mu \sum_i (c_i - \eta_i)' x_i\beta \right.\right.$$
$$-\beta' h_\mu \sum_i x_i'(c_i - \eta_i) + \beta' h_\mu \sum_i x_i'x_i\beta + \beta'\Sigma_0^{-1}\beta - \beta'\Sigma_0^{-1}B_0$$
$$\left.\left. -B_0'\Sigma_0^{-1}\beta + B_0'\Sigma_0^{-1}B_0 \right]\right\} \quad \text{(A.7)}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\beta'\left(h_\mu \sum_i x_i'x_i + \Sigma_0^{-1}\right)\beta - \beta'\left(h_\mu \sum_i x_i'(c_i - \eta_i) + \Sigma_0^{-1}B_0\right) \right.\right.$$
$$\left.\left. -\left(h_\mu \sum_i (c_i - \eta_i)' x_i + B_0'\Sigma_0^{-1}\right)\beta\right]\right\} \quad \text{(A.8)}$$

Calling $a = h_\mu \sum_i x_i'x_i + \Sigma_0^{-1}$ and $b = h_\mu \sum_i x_i'(c_i - \eta_i) + \Sigma_0^{-1}B_0$,

$$= \exp\left[-\frac{1}{2}\left(\beta'a\beta - \beta'b - b'\beta\right)\right]$$
$$= \exp\left[-\frac{1}{2}\left(\beta'a\beta - \beta'b - b'\beta + b'a^{-1}b - b'a^{-1}b\right)\right] \quad \text{(A.9)}$$
$$= \exp\left[-\frac{1}{2}\left(\beta'a\beta - \beta'aa^{-1}b - b'a^{-1}a\beta + b'a^{-1}aa^{-1}b\right)\right]$$

And considering,

$$B_1 = a^{-1} \qquad \therefore \qquad a = B_1^{-1} \qquad\qquad \mu_1 = a^{-1}b \qquad \therefore \qquad \mu_1' = b'a^{-1} \quad \text{(A.10)}$$

$$= \exp\left[-\frac{1}{2}\left(\beta'B_1^{-1}\beta - \beta'B_1^{-1}\mu_1 - \mu_1'B_1^{-1}\beta + \mu_1'B_1^{-1}\mu_1\right)\right]$$
$$= \exp\left[-\frac{1}{2}(\beta - \mu_1)' B_1^{-1} (\beta - \mu_1)\right] \qquad \text{(A.11)}$$

Hence, $\beta|y, c \sim \mathcal{N}(\mu_1, B_1)$; where

$$B_1 = \left[h_\mu \sum_i x_i'x_i + \Sigma_0^{-1}\right]^{-1} \qquad \text{and} \qquad \mu_1 = B_1\left[h_\mu \sum_i x_i'(c_i - \eta_i) + \Sigma_0^{-1}B_0\right]$$
$$\text{(A.12)}$$

A.2   Conditional Posterior Distribution for h_μ

$$\pi\left(h_\mu|y,\,c\right) = \mathcal{L}\left(y|\theta,\,\eta\right)\mathcal{L}\left(c|\theta,\,\eta\right)\pi\left(\eta|\theta\right)\pi\left(\theta_{-h_\mu}\right)\pi\left(h_\mu\right)$$

$$\text{(A.13)}$$

$$\propto \exp\left[\sum_i \nu_i\left(z_i'\gamma + \delta\eta_i\right)\right]\exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$h_\mu^{\frac{N}{2}}\exp\left[-\frac{1}{2}h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right)\right]$$

$$\pi\left(\eta|\theta\right)\pi\left(\theta_{-h_\mu}\right)\frac{\frac{v_\mu}{2}^{\frac{s_\mu}{2}}}{\Gamma\left(\frac{s_\mu}{2}\right)}h_\mu^{\frac{s_\mu}{2}-1}\exp\left(-\frac{v_\mu}{2}h_\mu\right) \quad \text{(A.14)}$$

$$\propto h_\mu^{\frac{N}{2}}\exp\left[-\frac{1}{2}h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right)\right]$$

$$h_\mu^{\frac{s_\mu}{2}-1}\exp\left(-\frac{v_\mu}{2}h_\mu\right) \quad \text{(A.15)}$$

$$= h_\mu^{\left(\frac{s_\mu}{2}+\frac{N}{2}\right)-1}\exp\left\{-\left[\frac{1}{2}\sum_i \left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right) + \frac{v_\mu}{2}\right]h_\mu\right\} \quad \text{(A.16)}$$

Hence, $h_\mu|y,\,c \sim \mathcal{G}\left(\frac{s_{1\mu}}{2},\,\frac{v_{1\mu}}{2}\right)$; where

$$s_{1\mu} = s_\mu + N \qquad \text{and} \qquad v_{1\mu} = \sum_i \left(c_i - x_i\beta\eta_i\right)'\left(c_i - x_i\beta - \eta_i\right) + v_\mu \quad \text{(A.17)}$$

A.3   Conditional Posterior Distribution for h_η

$$\pi\left(h_\eta|y,\,c\right) = \mathcal{L}\left(y|\theta,\,\eta\right)\mathcal{L}\left(c|\theta,\,\eta\right)\pi\left(\eta|\theta\right)\pi\left(\theta_{-h_\eta}\right)\pi\left(h_\eta\right)$$

$$\text{(A.18)}$$

$$\propto \exp\left[\sum_i \nu_i\left(z_i'\gamma + \delta\eta_i\right)\right]\exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$h_\mu^{\frac{N}{2}}\exp\left[-\frac{1}{2}h_\mu\left(c_i - x_i\beta - \eta_i\right)'\left(c_i - x_i\beta - \eta_i\right)\right]$$

$$h_\eta^{\frac{N}{2}}\exp\left[-\frac{1}{2}h_\eta \sum_i \eta_i^2\right]\pi\left(\theta_{-h_\eta}\right)\frac{\frac{v_\eta}{2}^{\frac{s_\eta}{2}}}{\Gamma\left(\frac{s_\eta}{2}\right)}h_\eta^{\frac{s_\eta}{2}-1}\exp\left(-\frac{v_\eta}{2}h_\eta\right) \quad \text{(A.19)}$$

29

$$\propto h_\eta^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\eta \sum_i \eta_i^2\right] h_\eta^{\frac{s_\eta}{2}-1} \exp\left(-\frac{v_\eta}{2} h_\eta\right)$$

$$= h_\eta^{\left(\frac{s_\eta}{2}+\frac{N}{2}\right)-1} \exp\left[-\left(\frac{1}{2}\sum_i \eta_i^2 + \frac{v_\eta}{2}\right) h_\eta\right] \tag{A.20}$$

Likewise, $h_\eta \,|\, y,\, c \sim \mathcal{G}\left(\frac{s_{1\eta}}{2},\, \frac{v_{1\eta}}{2}\right)$; where,

$$s_{1\eta} = s_\eta + N \qquad \text{and} \qquad v_{1\eta} = \sum_i \eta_i^2 + v_\eta \tag{A.21}$$

A.4   Conditional Posterior Distribution for γ

$$\pi\left(\gamma | y,\, c\right) = \mathcal{L}\left(y | \theta,\, \eta\right) \mathcal{L}\left(c | \theta,\, \eta\right) \pi\left(\eta | \theta\right) \pi\left(\theta_{-\gamma}\right) \pi\left(\gamma\right) \tag{A.22}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$h_\mu^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)' \left(c_i - x_i\beta - \eta_i\right)\right]$$

$$\pi\left(\eta | \theta\right) \pi\left(\theta_{-\gamma}\right) \frac{1}{(2\pi)^{\frac{j}{2}} |\Omega_0|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\left(\gamma - \Gamma_0\right)' \Omega_0^{-1}\left(\gamma - \Gamma_0\right)\right] \tag{A.23}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$\exp\left[-\frac{1}{2}\left(\gamma - \Gamma_0\right)' \Omega_0^{-1}\left(\gamma - \Gamma_0\right)\right] \tag{A.24}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right) - \sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right) - \frac{1}{2}\left(\gamma - \Gamma_0\right)' \Omega_0^{-1}\left(\gamma - \Gamma_0\right)\right] \tag{A.25}$$

A.5   Conditional Posterior Distribution for δ

$$\pi\left(\delta | y,\, c\right) = \mathcal{L}\left(y | \theta,\, \eta\right) \mathcal{L}\left(c | \theta,\, \eta\right) \pi\left(\eta | \theta\right) \pi\left(\theta_{-\delta}\right) \pi\left(\delta\right) \tag{A.26}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$h_\mu^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\mu \sum_i (c_i - x_i\beta - \eta_i)' (c_i - x_i\beta - \eta_i)\right]$$

$$\pi(\eta|\theta) \pi(\theta_{-\delta}) \frac{h_\delta^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}} \exp\left[-\frac{1}{2}h_\delta (\delta - \Delta_0)' (\delta - \Delta_0)\right] \quad \text{(A.27)}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$\exp\left[-\frac{1}{2}h_\delta (\delta - \Delta_0)' (\delta - \Delta_0)\right] \quad \text{(A.28)}$$

$$\propto \exp\left[\sum_i \nu_i (z_i'\gamma + \delta\eta_i) - \sum_i y_i \exp(z_i'\gamma + \delta\eta_i) - \frac{1}{2}h_\delta (\delta - \Delta_0)' (\delta - \Delta_0)\right] \quad \text{(A.29)}$$

A.6   Conditional Posterior Distribution for the Inefficiency (η_i)

$$\pi(\eta|y,\,c) = \mathcal{L}(y|\theta,\,\eta)\,\mathcal{L}(c|\theta,\,\eta)\,\pi(\eta|\theta)\,\pi(\theta)$$

$$\text{(A.30)}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$h_\mu^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\mu \sum_i (c_i - x_i\beta - \eta_i)' (c_i - x_i\beta - \eta_i)\right]$$

$$h_\eta^{\frac{N}{2}} \exp\left[-\frac{1}{2}h_\eta \sum_i \eta_i^2\right] \quad \text{(A.31)}$$

$$\propto \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right)\right] \exp\left[-\sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right]$$

$$\exp\left[-\frac{1}{2}h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)' \left(c_i - x_i\beta - \eta_i\right)\right]$$

$$\exp\left[-\frac{1}{2}h_\eta \sum_i \eta_i^2\right] \qquad \text{(A.32)}$$

$$= \exp\left[\sum_i \nu_i \left(z_i'\gamma + \delta\eta_i\right) - \sum_i y_i \exp\left(z_i'\gamma + \delta\eta_i\right)\right.$$

$$\left. -\frac{1}{2}h_\mu \sum_i \left(c_i - x_i\beta - \eta_i\right)' \left(c_i - x_i\beta - \eta_i\right) - \frac{1}{2}h_\eta \sum_i \eta_i^2\right] \quad \text{(A.33)}$$

B   Robustness Check

In order to check the robustness of our results (see Table 4.1) we repeated the estimations using a different set of covariates as well as the average of the last available financial statements prior to the bank failure. All results are similar to the ones presented in Table 4.1. As an example, we present Table B.1 below.

Table B.1: **Estimation results (Robustness Checks)**

	Bayesian estimation				ML estimation		
	Mean	Perc. 2.5$^{\text{th}}$	Perc. 50$^{\text{th}}$	Perc. 97.5$^{\text{th}}$	Estimates	t-test	
$Cons$	0.3893	0.0480	0.3895	0.7344	1.3460	2.4462	*
Ln(y_1)	0.1735	0.1456	0.1735	0.1997	0.1760	3.4368	*
Ln(y_2)	0.8809	0.8479	0.8807	0.9148	0.8205	11.1307	*
Ln(y_1)Ln(y_2)	-0.0921	-0.0935	-0.0921	-0.0907	-0.0976	-54.0006	*
Ln(w_1/w_3)	-0.2319	-0.2653	-0.2317	-0.1980	-0.2371	-1.7811	
Ln(w_2/w_3)	0.0604	-0.0094	0.0612	0.1300	-0.0658	-0.3645	
Ln(w_1/w_3)Ln(w_2/w_3)	0.0487	0.0445	0.0488	0.0529	0.0465	2.5263	*
Ln(y_1)^2	0.0806	0.0792	0.0806	0.0820	0.0853	24.7009	*
Ln(y_2)^2	0.0811	0.0787	0.0811	0.0835	0.0853	11.0330	*
Ln(w_1/w_3)^2	-0.0372	-0.0407	-0.0372	-0.0337	-0.0301	-2.3634	*
Ln(w_2/w_3)^2	0.0049	-0.0036	0.0050	0.0135	0.0061	0.1823	
Ln(y_1)Ln(w_1/w_3)	-0.0416	-0.0431	-0.0416	-0.0398	-0.0519	-9.1977	*
Ln(y_1)Ln(w_2/w_3)	0.0444	0.0414	0.0444	0.0475	0.0494	4.6703	*
Ln(y_2)Ln(w_1/w_3)	0.0331	0.0311	0.0331	0.0351	0.0426	7.0642	*
Ln(y_2)Ln(w_2/w_3)	-0.0354	-0.0390	-0.0355	-0.0317	-0.0308	NaN	
σ_μ	0.0460	0.0437	0.0460	0.0484	0.1287	NaN	
σ_η	0.4896	0.4816	0.4896	0.4975	0.7224	NaN	
Camel C_1	-1.4387	-2.7038	-1.6838	-0.1259	-1.2844	-0.8665	
Camel A_1	1.6764	0.7607	1.7635	2.1717	-1.7637	-2.3136	*
Camel A_2	1.6467	0.1833	1.6209	2.6327	1.6350	3.4616	*
Camel A_4	0.2132	-0.1200	0.1899	0.6020	-0.8300	-0.0163	
Camel A_6	0.3184	-0.1280	0.3522	0.7229	2.3650	0.0331	
Camel E_1	-1.1971	-2.1794	-1.1766	-0.2410	-1.2652	-1.6595	
Camel L_1	0.4915	0.0214	0.5126	0.9372	3.2086	12.8973	*
Ln $Assets$	-0.7366	-0.8263	-0.7511	-0.5969	-0.7489	-13.8182	*
Ineff δ	0.5966	0.2664	0.5971	0.9510	1.0178	3.7896	*

Note: This table shows the results of the estimation by the proposed simultaneous Bayesian method and the two-stage approach by maximum likelihood. The estimates with an asterisk are statistically significant at 5% confidence level. Our parameter of interest (δ) is significant and it has the expected positive sign. It means that being more inefficient increases the failure function rate and consequently decreases the mean time to failure. That is, it increases the probability of failure.