Journal of Statistical and Econometric Methods, vol.3, no.1, 2014, 153-178
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2014

An Application of Robust Regression to Bernanke's Analysis of Nonmonetary Effects in the Great Depression

Christopher V. Rackauckas[1]

Abstract

Bernanke's "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression" has been influential in the macroeconomic community by creating a study of the nonmonetary effects of the financial markets on macroeconomic activity. In this work, he hypothesized that the weakening of the financial system leads to an economic contraction through an additional nonmonetary factor. However, the data set utilized for this study had a large outlier corresponding to the bank holiday in March 1933. We see that omitting the outlier leads to results that do support Bernanke's hypothesis. However, Bernanke argues that the outlier cannot be simply omitted as it holds valuable information about the chaotic state of the financial markets in that time period. Thus we used robust statistics to incorporate the effect outliers in a purely statistical manner. The result shows that nonmonetary effects from the financial markets are indeed significant according to the robust estimators supporting Bernanke's hypothesis.

[1] Department of Mathematics, University of California, Irvine,
e-mail: Contact@ChrisRackauckas.com

Article Info: *Received* : December 9, 2013. *Revised* : January 25, 2014.
Published online : February 7, 2014

Mathematics Subject Classification: 62F35; 62F15; 62J05; 91B84
Keywords: Robust Multiple Linear Regression Regression; Bayesian Estimation; M-Estimators; Nonmonetary Effects

1 Introduction

In 1983 the now Chairman of the Federal Reserve Ben Bernanke published a paper titled *Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression* empirically analyzing the effects of credit channels on macroeconomic activity [1]. However, the effects of the bank holiday on March 1933 are represented in the data set as a large outlier in one of the explanatory series which he argues could not be completely eliminated due to the nature of the event. In order to deal with such a necessary outlier, Bernanke implements a strong assumption to scale down the outlier's effect in an ordinary least square (OLS) regression. The goal of this paper is to investigate the results of various mathematical techniques that could have been used to make the regression more robust to outliers. Specifically, Bayesian Robust Regression and M-estimator methods are introduced and applied to the data set.

The outline of this paper is as follows. Section 2 outlines the economic theory presented in Bernanke's paper to give the reader background in the economic theory. Section 3 outlines the mathematical theory behind the robust regression techniques of interest. Section 4 explains the estimation methodology and compares the results of the robust regressions to the results of the OLS regression employed by Bernanke.

2 Nonmonetary Effects in the Great Depression

Bernanke's *Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression* has been influential in the macroeconomic community and has stirred much interest in the empirical evaluation of the impact of credit channels on macroeconomic activity [2, 3]. In this work he points out that financial crises coincide with adverse developments in the macroeconomy and argues that it must be explained by a causal relation between these entities. Bernanke discounts the idea that movements in the financial system simply responded without feedback to declines in aggregate output since problems in the financial system tend to lead output declines. Instead, he adopts the position of Friedman and Schwartz who argue that a weakening of the financial system causes an economic contraction by reducing the wealth of bank shareholders and causing a rapid decrease in the supply of money.

Bernanke adds that the weakening of the financial system leads to an economic contraction through an additional factor, an intermediation between borrowers and lenders due to the requirement of a nontrivial market-making and information-gathering service in the real financial market. He argues that if we assume that information and transaction costs exist within financial markets, then there exists a nontrivial cost of credit intermediation (CCI) for lenders to provide the necessary screening, monitoring, and accounting for a bank to successfully minimize the number of loans given to "bad borrowers" (this cost would also include the expected losses inflicted by lending to bad borrowers). Banking crises and the prevalency of bankruptcies are argued as negatively affecting the CCI by causing the banks to be more careful in choosing borrowers which in turn leads to higher degrees of borrower screening, monitoring, and accounting. Bernanke relates the CCI to aggregate demand by noting that a higher CCI implies that borrowers face a higher cost of credit (but no higher cost of saving) which reduces demand for current-period goods. This downward shift of the demand curve implies a decrease in aggregate output and thus the CCI is negatively correlated to aggregate output. The important fact to note is that this argument shows that there exists a nonmonetary effect of the financial crisis, the the effect due to the CCI, on output.

To estimate the effect of the CCI on output, Bernanke decided to empirically analyze the CCI relative to industrial production in the early events of the Great Depression. Since no series for the CCI is readily available, Bernanke used the deposits of failing banks and the liabilities of failing commercial businesses as proxies. The reasoning for the deposits of failing banks is that increases in deposits of failing banks are correlated with the CCI since it would imply that banks are more stringent on giving loans meaning that finding appropriate borrowers is a more costly task. As the liabilities of failing commercial businesses increases, the number of businesses that are seen as potentially good borrowers decreases and thus the CCI would increase to reflect the increased difficulty of finding an appropriate borrower.

However, the national bank holiday in March 1933 lead to the deposits of banks to be seven times worse than that in the deposits of failing banks. This large figure reflects the chaotic financial conditions of the time and resembles earlier crises in such a manner that simply throwing away the point would be discarding of actual information. To deal with the inclusion of this outlier, Bernanke reasons that the closure of banks by government action would have created less fear than a similar response without government intervention. Thus he assumes that "supervised" bank closings in March 1933 had the same effects as an "unsupervised" bank crisis involving 15 percent of the frozen deposits. In effect, this allows Bernanke to scale down the March 1933 event to around the size of the events of October 1931, the next largest data point. From his empirical estimations using this scaled down version of the outlier, Bernanke finds that even when controlling for the effects of money supply and demand, the effects of the financial market are still significant and thus show a nonmonetary effect (presumably due to some effect like the CCI).

3 Robust Regression

Robust statistics are a set of theories and techniques for estimating the parameters of models while dealing with deviations from the assumptions commonly placed on these models (an introduction to the topic can be found at [4]). One such deviation is the contamination of data by gross errors (more commonly referred to as outliers). Gross errors can be defined as "data severely deviating from the pattern set by the majority of the data". Such an error can mean that the real data may not be normally distributed, which in turn could spoil the OLS estimates. The theory of robust regressions was developed to estimate the regression coefficients in a manner that incorporates gross errors. The techniques that we will be looking at are a technique for implementing robustness into a Bayesian multiple regression model and the most commonly used frequentest robust regression technique, M-estimation, using the Huber estimator.

3.1 Bayesian Robust Regression

The standard technique for multiple regression in a Bayesian context is as follows [5, 6, 7, 8, 9]. We assume that our data set is of the form $\{X_i, Y_i\}_{i=1}^N$ where X_i is a $K \times 1$ vector of predictor variables. Let $\beta = (\beta_1, \ldots, \beta_k)'$ a vector of random variables representing the unknown parameters. The model that we are estimating, like in the frequentest regression, is

$$Y_i = \beta X_i + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma^2)$ where σ^2 is a random variable representing an unknown parameter. In order to run the procedure, the user must specify a posterior distribution on the random variables. A normal distribution is commonly used for the parameters,

$$\beta_i \sim N(\mu_i, \sigma_i^2)$$

where μ_i is the prior "best guess" for β_i and σ_i^2 is the confidence in the guess (with large σ_i^2 indicating low confidence or a "weak/non-informative prior"). For σ^2, it is common to use a uniform distribution for the prior

$$\sigma^2 \sim U(0, 10)$$

since σ^2 will fall into this range for almost any data set[2]. The likelihood is specified as

$$Y_i \sim N(\beta X_i, \sigma^2).$$

A hierarchical diagram of this setup is shown in Figure 1.The Bayesian method for computing the best estimate for distributions of the random variables when incorporating the data uses Bayes Theorem. The best estimate for the distributions can be found by the equation

$$Posterior(\beta) \propto Likelihood(\beta) \times Prior(\beta).$$

This can be solved computationally using Markov Chain Monte Carlo (MCMC) and Gibbs Sampling techniques employed in software packages such as JAGS

[2]This is because the data are standardized before the computation is done in order to lead to faster convergence of the Gibbs sampling method. Thus data distributions close to normal, over 99% of the data will fall within 3 of the mean. Thus by allowing σ^2 to go up to 10, any value of σ^2 that would be required for a standard data set is a possible outcome. If σ^2 is actually above 10, then the posterior distribution would look "squished" towards 10, the highest non-zero value of the prior.

(Just Another Gibbs Sampler). The estimators for the regression, $\hat{\beta}$, are the means of the posterior distributions of the parameters. It can be shown that when using non-informative priors the results of this analysis converge to the OLS estimators [6].

The Bayesian multiple regression can be made robust to outliers by changing the distribution of the likelihood function to be a Student's t-distribution as

$$Y_i \sim t(\beta X_i, \sigma^2, \nu)$$

where ν is a random variable for the degrees of freedom of the t-distribution. The degrees of freedom can be set by transforming it to another random variable u where

$$\nu = 1 - k * \log(1 - u)$$

with some constant k which expresses a prior belief in the value of ν (large k implies a prior belief of a large ν). A uniform distribution can be placed on u,

$$u \sim Unif(0, 1)$$

which places the distribution of ν in the interval $[1, \infty)$.The t-distribution is a symmetric distribution which has fatter tails than the normal distribution. The idea is that these fatter tails are able to incorporate gross errors without shifting the mean of the distribution as much as would occur if the distribution was a normal distribution. The posterior estimates from the model will choose the most appropriate distribution for the degrees of freedom given the data and estimate the regression coefficients in a manner that is less pulled by gross errors. Since as the degrees of freedom in the t-distribution approaches infinity the distribution approaches a normal distribution, we can see the "normality" of our model by looking at the mean of the posterior distribution for ν.

The Bayesian version of hypothesis testing instead uses highest density intervals (HDI). The 95% HDI is the interval of the posterior distribution for a random variable which holds the parts of the distribution with the highest density and integrates to give .95. Thus it is the 95% interval where the parameter is most likely to exist. The hypothesis tests for Bayesian statistics then uses the HDI as a substitute for the confidence intervals and the same rules for testing a null hypothesis apply. However, Bayesian 95% HDI's generated from a multiple regression correspond to the confidence intervals from a Tukey

Honestly Significant Difference test and thus problems with type-1 error due to multiple testing do not apply in the Bayesian context. Thus inference can be made from the posterior HDIs without worry of multiple comparison problems.

3.2 M-Estimation Using the Huber Estimator

The most common method of robust regression is M-estimation [10]. Consider the linear model $Y_i = X_i\beta + \epsilon_i$. The fitted model can be written as $Y_i = X_i\hat{\beta} + e_i$. The general M-estimator minimizes the objective function

$$\sum_{i=1}^{n} p(e_i) = \sum_{i=1}^{n} p(y_i - X_i\hat{\beta})$$

where the function p gives the contribution of each residual to the objective function. p must uphold the following properties:

1. $p(e) \geq 0$
2. $p(0) = 0$
3. $p(e) = p(-e)$
4. $p(e_i) \geq p(e_{i'})$ for $|e_i| > |e_{i'}|$

For example, the ordinary least squares estimation uses $p(e_i) = e_i^2$. Optimization techniques can be used to solve for the $\hat{\beta}$ that best solves this equation.

Robust regression is implemented by setting the p function to the Hubor p defined as

$$p(e_i) = \begin{cases} \frac{1}{2}e^2 & |e| \leq k \\ k|e| - \frac{1}{2}k^2 & |e| > k \end{cases}$$

where k is a tuning constant that gives more resistance to outliers with lower values. To give high efficiency in the estimation, k is usually chosen as $k = 1.345\sigma$. This equation could then be solved using a maximum likelihood estimation through a optimization methods such as iteratively reweighted least-squares (IRLS). The result gives the maximum likelihood estimates and their associated t-statistics. However, it should be noted that these t-statistics do not necessarily t-distributed. Thus the standard t-test is not application. Either distribution theory or bootstrapping techniques must be applied to the t-statistics to find appropriate p-values, though the same idea would apply that larger t-statistics are associated with small p-values.

4 Empirical Estimation

We will use the procedure laid out by Bernanke in order to estimate the nonmonetary effects on output. Along the way we will compare our results to those of both Bernanke and the work of Miron and Rigol who replicated the study in 2012 [2]. The procedure is as follows. We will attempt to use the same data sets as those Bernanke used. The data set starts on January 1919 and goes until December 1941. For output we will use the industrial production index from the *Federal Reserve Bulletin* adjusted to have $01/1930 = 100$. Money supply will be measured using M1 from the series in Friedman and Schwartz [11]. Our proxy for money demand will be prices will be measured using the wholesale price index adjusted to have 1957-1959=100. As noted before, we will use the deposits of failing banks and the liabilities of failing commercial businesses as proxies for the effect of the CCI[3]. The series for the deposits of failing banks comes from the *Federal Reserve Bulletin.* The series for the liabilities of failing banks comes from Dun's series.

Shocks to money supply should be associated with changes in output. The series of interest is the rate of growth of output relative to exponential trend denoted by y_t. The monetary effects on output are from shocks to money supply and money demand. The shocks to money supply, that is the money supply minus the expected money supply, $(M-M^e)_t$, are defined as the residuals from a regression of the rate of growth of M1 on four lags of the growth rates of output (y), prices, and M1. The shocks to money demand, price minus the expected price, $(P-P^e)_t$ are defined symmetrically. We assume that shocks to output would be caused by changes in the deposits of failing banks and the

[3]Note that Bernanke and Miron and Rigol use the deposits of failing banks in millions whereas the data series I employ uses the deposits in thousands. This leads to the appropriate scaling factor for the comparison between our study and theirs as 1000 for the coefficient of variables for the deposits of failing banks.

liabilities of failing commercial businesses, and thus we define $DBanks$ and $DFail$ as the first difference of their respective series deflated by the wholesale price index. To determine the effect of money supply on output, the regression equation

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 (M - M^e)_t + \beta_4 (M - M^e)_{t-1} + \beta_5 (M - M^e)_{t-2} + \beta_6 (M - M^e)_{t-3} \tag{1}$$

is estimated. Likewise, to determine the effect of money demand on output, the regression equation

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 (P - P^e)_t + \beta_4 (P - P^e)_{t-1} + \beta_5 (P - P^e)_{t-2} + \beta_6 (P - P^e)_{t-3} \tag{2}$$

is estimated. To test the hypothesis that Bernanke proposed, we wish to estimate a regression equation which controls for money supply and demand and tells us the effect of bank and business failures on output in a nonmonetary way. To do so, we estimate the regression equations

$$\begin{aligned} y_t \;=\; & \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 (M - M^e)_t + \beta_4 (M - M^e)_{t-1} + \beta_5 (M - M^e)_{t-2} \qquad (3) \\ & + \beta_6 (M - M^e)_{t-3} + \beta_7 DBanks_t + \beta_8 DBanks_{t-1} + \beta_9 DFail_t + \beta_{10} DFail_{t-1} \end{aligned}$$

$$\begin{aligned} y_t \;=\; & \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 (P - P^e)_t + \beta_4 (P - P^e)_{t-1} + \beta_5 (P - P^e)_{t-2} \qquad (4) \\ & + \beta_6 (P - P^e)_{t-3} + \beta_7 DBanks_t + \beta_8 DBanks_{t-1} + \beta_9 DFail_t + \beta_{10} DFail_{t-1} \end{aligned}$$

4.1 Ordinary Least Squares Replication

The results of these regressions using OLS are shown in Table 1. The tables correspond to the regression equations shown above. Also listed are the results of Miron and Rigol's replication and the results of a Bayesian estimation using a normal likelihood (non-robust Bayesian Multiple Linear Regression) and a non-informative prior of $\beta_i \sim N(0, 1 \times 10^{-12})$ and $\sigma^2 \sim U(0, 10)$. For the Bayesian method, the means of the distribution are reported in the table and graphs of their distributions can be found as figures 2 through 5. Notice that our results agree closely with the replication of Miron and Rigol's. The coefficients of the shocks to money and price in equations one and two are positive and significant which is consistent with the Friedman and Schwartz view that money shocks were an important aspect in the decline of output during the Great Depression. The lags on the shocks are not significant which is consistent with a rational expectations framework in that only unpredicted shocks to money should affect real variables [2]. Lastly, the Bayesian coefficients closely resemble those of the OLS coefficients which is what was expected given our non-informative prior and normal likelihood [8].

The coefficients on the terms for the liabilities of failed banks are both negative and significant. This matches what we expected from the theory presented in Section 2 which suggest that even after controlling for the impact of bank failures on the money stock, there would be an additional non-monetary effect (the CCI for example) leading to declines in output. The coefficients of business liabilities also give negative results which further Bernanke's hypothesis.

4.2 Outlier Results and Robust Estimations

The bank holiday in March of 1933 presents a problem for our data set. The deposits of banks suspended in March 1933 is seven times as large as the next highest reading. This raises a question as to how much the data set is biased towards this reading. Bernanke argues that it would be a mistake leave out the bank holiday since it was a response to the panicky financial conditions of the period. Table 2 shows the OLS regression results excluding the outlier from the bank holiday. Also listed are the results from Miron and Rigol who also ran the regression omitting the bank holiday. Notice that by doing so, the significance of the deposits of failing banks and the business liabilities is lost and thus provide evidence against Bernanke's hypothesis.

This shows that this data point not only has a large effect, but the strength of the effect determines whether or not the coefficients for the deposits of failing banks and the business liabilities, the terms that lead to the confirmation of Bernanke's hypothesis, are significant. To account for this, Bernanke scaled down the size of the March 1933 reading by 15% to about the size of the event on October 1931 (claiming that this would be about the effect of an "unsupervised" bank crisis). Instead, we turn to the robust statistical methods discussed in Section 3. Table 3 shows Bernanke's results and the results, Miron and Rigol's results from scaling the outlier in the same manner as Bernanke, the results from the robust statistics. For the Bayesian method, the means of the distribution are reported in the table and graphs of their distributions can be found as figures 6 through 9. The results from the M-estimator are not as readily tested since the t-statistics do not follow a t-distribution and thus some technique such as bootstrapping would have to be employed in order to receive the p-values for the t-statistics. However, large t-statistics still correspond to small p-values and thus the large t-statistics on $DBanks$ in the estimation of equations 3 and 4 provide evidence for Bernanke's hypothesis. The Bayesian robust estimates provide the least variance on the likelihood/error terms, indicating the best fit of all of the models. Here we see that when controlling for the monetary and price effects find an negative effect due to $DBanks$ which supports Bernanke's hypothesis. Notice that when using the Bayesian robust estimates only price shocks has significant first lags. The lack of significance on the lags makes the model more consistent with a rational expectations framework than the non-robust estimates. Notice that that degrees of freedom all fall within 3 and 4, indicating that the distribution of the data is not normal and is better modeled by a t with low degrees of freedom.

The large decrease in the standard error and the low degrees of freedom in the estimated likelihood t-distribution show that a robust statistical method such as the Bayesian Robust Multiple Linear Regression employed here provides us with better estimators than OLS. The estimations that we receive when performing such a robust regression not only support Bernanke's hypothesis that even when controlling for the monetary effects of bank failures there exists a nonmonetary effect on output, but because the robust estimates drop the significance of the first lag of $DBanks$, it fits a rational expectations framework better than the non-robust estimates. One thing to note is that the effect of the liabilities of failing banks was not able to be seen in the robust regression. This effect was fairly weak in our non-robust estimations. A reason this may be occurring is the fact that the series was heavily revised in 1933 "to exclude real estate and insurance brokers, holding and finance companies, shipping agents, tourist companies, transportation terminals, and such". This large break in 1933 may have damaged some of the estimates of this series. However, overall it seems that the robust regression methods provide a confirmation of Bernanke's hypothesis.

5 Conclusion

Bernanke's *Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression* has been influential in the macroeconomic community by creating a study of the nonmonetary effects of the financial markets on macroeconomic activity. However, the data set utilized for this study had a large outlier corresponding to the bank holiday in March 1933. We see that omitting the outlier leads to results that do not support Bernanke's hypothesis. However, Bernanke argues that the outlier cannot be simply omitted as it holds valuable information about the chaotic state of the financial markets in that time period. Thus we used robust statistics to incorporate the effect outliers in a purely statistical manner. The result shows that nonmonetary effects from the financial markets are indeed significant according to the robust estimators, supporting Bernanke's hypothesis.

Appendix

Bayesian Robust Multiple Linear Regression Software

The software for performing the Bayesian Robust Multiple Linear Regressions can be found at www.chrisrackauckas.com. It is an R program developed by Chris Rackauckas and is modified from the Bayesian Multiple Linear Regression package by John K. Kruschke. The program utilizes the rjags package to interface with the JAGS Gibss Sampler to perform the Monte Carlo Markov Chain numerical estimations of the posterior distributions.

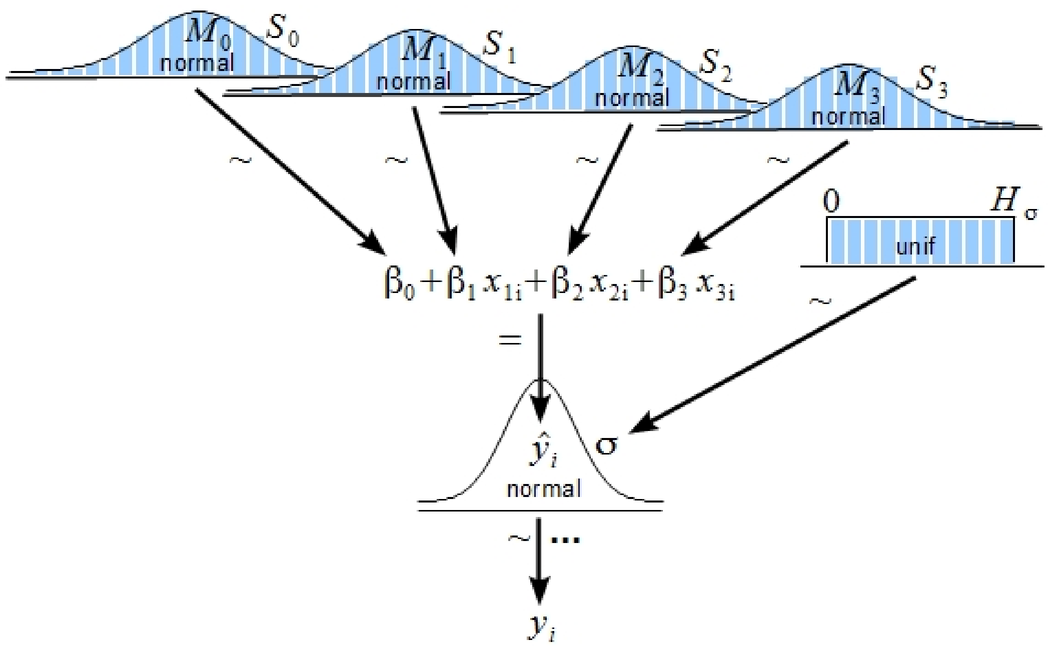

Figure 1: Hierarchical diagram of a multiple linear regression model with three predictor variables. The prior distribution is shown with histogram bars superimposed to indicate the correspondence with the posterior distributions shown in subsequent figures [8].

Table 1: OLS Regression Results

Coe cient	OLS Eq1	OLS Eq2	OLS Eq3	OLS Eq4	M&R Eq1	M&R Eq2	M&R Eq3	M&R Eq4	Bayes Eq1	Bayes Eq2	Bayes Eq3	Bayes Eq4
y_{t-1}	0.624***	0.535***	0.609***	0.595***	0.669***	0.601***	0.654***	0.619***	0.624***	.0535***	.619***	0.595***
	(10.12)	(8.70)	(9.82)	(9.50)	(10.20)	(9.16)	(10.06)	(9.54)				
y_{t-2}	-0.0857	-0.0542	-0.109	-0.0777	-0.117	-0.105	-0.131*	-0.102	-0.0857	-.0541	-.110	-0.0776
	(-1.40)	(-0.923)	(-1.77)	(-1.29)	(-1.78)	(-1.64)	(-2.05)	(-1.62)				
$(M-M^e)_t$	0.684***		0.396*		0.644***		0.308		0.684***		.395*	
	(4.67)		(2.495)		(4.40)		(1.90)					
$(M-M^e)_{t-1}$	0.0637		-0.0104		-0.078		-0.104		0.0641		-0.0109	
	(0.42)		(-0.07)		(-0.51)		(-0.67)					
$(M-M^e)_{t-2}$	-0.186		0.00762		-0.290		-0.055		-0.186		0.00764	
	(-1.25)		(0.05)		(-1.93)		(-0.35)					
$(M-M^e)_{t-3}$	-0.0782		-0.0484		-0.098		-0.071		-0.0782		0.00474	
	(-0.53)		(-0.33)		(-0.65)		(-0.49)					
$(P-P^e)_t$		0.599***		0.616***		0.784***		0.697***		0.599***		0.616***
		(5.33)		(4.69)		(5.50)		(5.09)				
$(P-P^e)_{t-1}$		0.523***		0.359**		.376*		0.248		0.524***		0.359**
		(4.42)		(2.65)		(2.47)		(1.71)				
$(P-P^e)_{t-2}$		0.1921		0.0588		0.140		0.013		0.193		0.0596
		(1.58)		(0.43)		(0.92)		(0.09)				
$(P-P^e)_{t-3}$		-0.149		-0.101		-0.124		-0.143		-0.149		-0.101
		(-1.23)		(-0.77)		(-0.83)		(-1.01)				
$DBanks_t$			-0.00000107***	-0.00000119***			-0.001047***	-0.001167***			-0.00000107***	-0.00000119***
			(-4.44)	(-6.17)			(-4.26)	(-6.04)				
$DBanks_{t-1}$			-0.000000528*	-0.000000571**			-0.000491	-0.000531**			-0.000000528*	-0.000000571**
			(-2.16)	(-2.82)			(-1.97)	(-2.62)				
$DFails_t$			-0.0160*	-0.0118			-0.01534*	-0.01013			-0.0160*	-0.0117
			(-2.40)	(-1.80)			(-2.30)	(-1.58)				
$DFails_{t-1}$			-0.0152*	-0.00974			-0.01356*	-0.00806			-0.00152*	-0.00971
			(-2.27)	(-1.49)			(-2.02)	(-1.25)				
Constant	-0.000431	0.00175	0.000609	0.00148	0.00177	0.00203	0.00184	0.00191	-0.000430	0.00175	0.000610	0.00148
	(-0.25)	(1.14)	(0.39)	(1.02)	(1.08)	(1.25)	(1.16)	(1.26)				
Standard Error	0.0259	0.0247	0.0237	0.0227	.0250	.02444	.0240	0.0228	0.0261	0.0248	0.0238	0.0228
N	261	261	239	239	234	234	234	234	261	261	239	239

*p<.05, **p<.01, ***p<.001. T-statistic is shown in parentheses. Bayesian significance levels are found by checking if 0 is inside the appropriate HDI.

Table 2: OLS Regression Results, March 1933 omitted.

Coe cient	OLS Eq1	OLS Eq2	OLS Eq3	OLS Eq4	M&R Eq1	M&R Eq2	M&R Eq3	M&R Eq4
y_{t-1}	.575***	.547***	0.567***	.592***	0.565***	0.542***	0.551***	0.530***
	(9.58)	(9.05)	(9.09)	(9.42)	(8.43)	(8.02)	(8.19)	(7.80)
y_{t-2}	-0.0764	-0.0398	-0.0827	-0.0682	-0.053	-0.006	-0.045	-0.002
	(-1.31)	(-0.69)	(-1.37)	(-1.13)	(-0.80)	(-0.09)	(-0.68)	(0.03)
$(M - M^e)_t$	0.488**		0.482**		0.435**		0.359*	
	(3.13)		(3.09)		(2.76)		(2.22)	
$(M - M^e)_{t-1}$	0.333*		0.301		0.082		0.036	
	(2.03)		(1.82)		(0.49)		(0.21)	
$(M - M^e)_{t-2}$	0.335*		0.275		0.208*		0.221	
	(2.04)		(1.65)		(1.26)		(1.34)	
$(M - M^e)_{t-3}$	0.0789		0.0179		0.207		0.223	
	(0.54)		(0.13)		(1.26)		(1.34)	
$(P - P^e)_t$		0.550***		0.567***		0.633***		0.585***
		(5.08)		(4.19)		(4.44)		(4.04)
$(P - P^e)_{t-1}$		0.455***		0.333*		0.244		0.216
		(4.03)		(2.44)		(1.63)		(1.42)
$(P - P^e)_{t-2}$		0.105		0.0504		-0.053		-0.046
		(0.91)		(0.37)		(-0.35)		(-0.30)
$(P - P^e)_{t-3}$		-0.156		-0.0913		-0.089		-0.101
		(-1.36)		(-0.69)		(-0.60)		(-0.68)
$DBanks_t$			-0.00000207	-0.00000134			-0.002342	-0.001478
			(-1.46)	(-0.96)			(-1.66)	(-1.08)
$DBanks_{t-1}$			-0.000000800	-0.000000592			-0.001275	-0.000808
			(-0.56)	(-0.42)			(-0.90)	(-0.59)
$DFails_t$			-0.00985	-0.0104			-0.00736	-0.00846
			(-1.44)	(-1.52)			(-1.07)	(-1.28)
$DFails_{t-1}$			-0.00980	-0.00843			-0.00946	-0.00872
			(-1.42)	(-1.23)			(-1.37)	(-1.32)
Constant	-0.00272	0.00120	-0.00153	0.00118	0.00179	0.00176	0.00191	0.00187
	(-1.65)	(0.82)	(-0.94)	(0.81)	(1.17)	(1.18)	(1.25)	(1.26)
Standard Error	0.0243	0.0233	0.0230	0.0227	0.0228	0.0222	0.0227	0.0222
N	258	258	236	236	225	225	225	225

*p<.05, **p<.01, ***p<.001. T-statistic is shown in parentheses.

Table 3: Robust Statistics Results.

Coe cient	Bern Eq1	Bern Eq2	Bern Eq3	Bern Eq4	MEst Eq1	MEst Eq2	MEst Eq3	MEst Eq4	Bayes Eq1	Bayes Eq2	Bayes Eq3	Bayes Eq4
y_{t-1}	0.623*** (10.21)	0.583*** (9.50)	0.613*** (10.21)	0.615*** (9.76)	0.648 (12.55)	0.582 (11.15)	0.645 (11.56)	0.640 (11.55)	0.657***	0.598***	.639***	0.648***
y_{t-2}	-.144* (-2.37)	-0.118 (-1.76)	-0.159*** (-2.63)	-0.131* (-2.13)	-0.098 (-1.91)	-0.0433 (-0.87)	-0.115 (-2.11)	-0.0792 (-1.49)	-0.104	-0.0425	-0.109	-0.0704
$(M-M^e)_t$	0.407** (3.42)		0.332** (2.92)		0.615 (5.00)		0.410 (2.92)		0.601***		0.471**	
$(M-M^e)_{t-1}$	1.41 (1.16)		0.113 (0.99)		0.197 (1.55)		0.0273 (0.20)		0.242		0.133	
$(M-M^e)_{t-2}$	0.513 (0.42)		0.110 (0.96		0.0114 (0.09)		-0.00629 (0.05)		0.0944		0.0897	
$(M-M^e)_{t-3}$	0.144 (1.19)		0.156 (1.38)		0.0143 (0.11)		-0.00803 (-0.06)		0.0684		0.0590	
$(P-P^e)_t$		0.533*** (5.33)		0.455*** (3.99)		0.527 (5.53)		0.618 (5.32)		0.521***		0.604***
$(P-P^e)_{t-1}$		0.350** (3.33)		0.231* (1.97)		0.551 (5.49)		0.491 (4.09)		0.560***		0.523***
$(P-P^e)_{t-2}$		0.036 (0.34)		-0.004 (-0.03)		0.0809 (0.78)		0.0562 (0.47)		0.0982		0.0963
$(P-P^e)_{t-3}$		0.069 (0.66)		0.024 (0.22)		-0.144 (-1.40)		-0.0900 (-0.78)		-0.136		-0.0763
$DBanks_t$			-0.000869*** (-4.24)	-0.000799*** (-4.03)			-0.00000145 (-6.81)	-0.00000116 (-6.75)			-0.00000142**	-0.00000113***
$DBanks_{t-1}$			-0.000406 (-1.93)	-0.000337 (-1.66)			-0.000000699 (-3.24)	-0.000000531 (-2.96)			-0.000000700	-0.000000526
$DFails_t$			-0.00258* (-1.95)	-0.00202 (-1.52)			-0.0108 (-1.82)	-0.00836 (-1.44)			-0.00584	-0.00619
$DFails_{t-1}$			-0.00325** (-2.47)	-0.00242 (-1.83)			-0.0140 (-2.36)	-0.00704 (-1.21)			-0.0113	-0.00567
Constant							0.000142 (0.10)	0.00185 (1.44)	-0.00221	0.00184	-0.000626	0.00197
ν									3.22	3.33	3.30	3.64
Standard Error	0.0272	0.0260	0.0249	0.0246	.0200	0.0192	0.0190	0.0192	0.0178	0.0171	0.01668	1.64
N					261	261	239	239	261	261	239	239

*p<.05, **p<.01, ***p<.001. T-statistic is shown in parentheses. Bayesian significance levels are found by checking if 0 is inside the appropriate HDI.

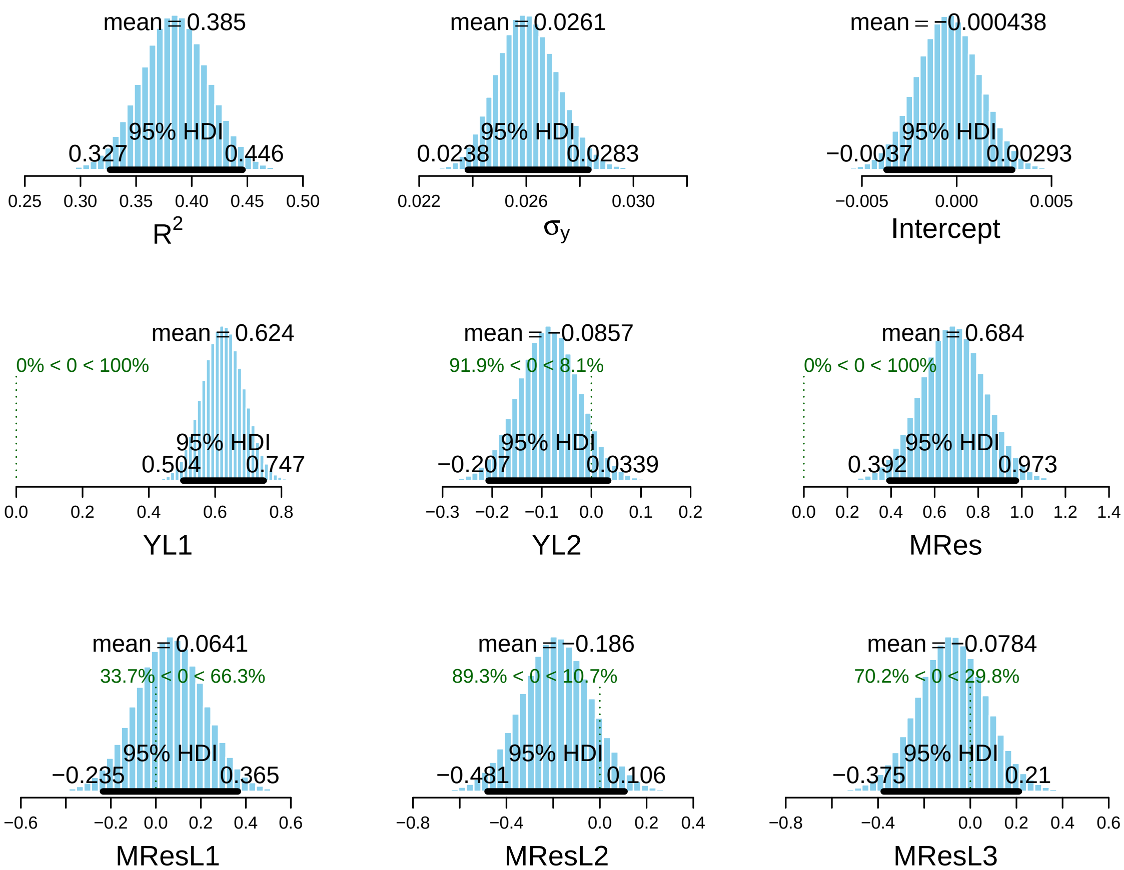

Figure 2: Posterior distributions of the parameters in Equation 1 using a normal likelihood.

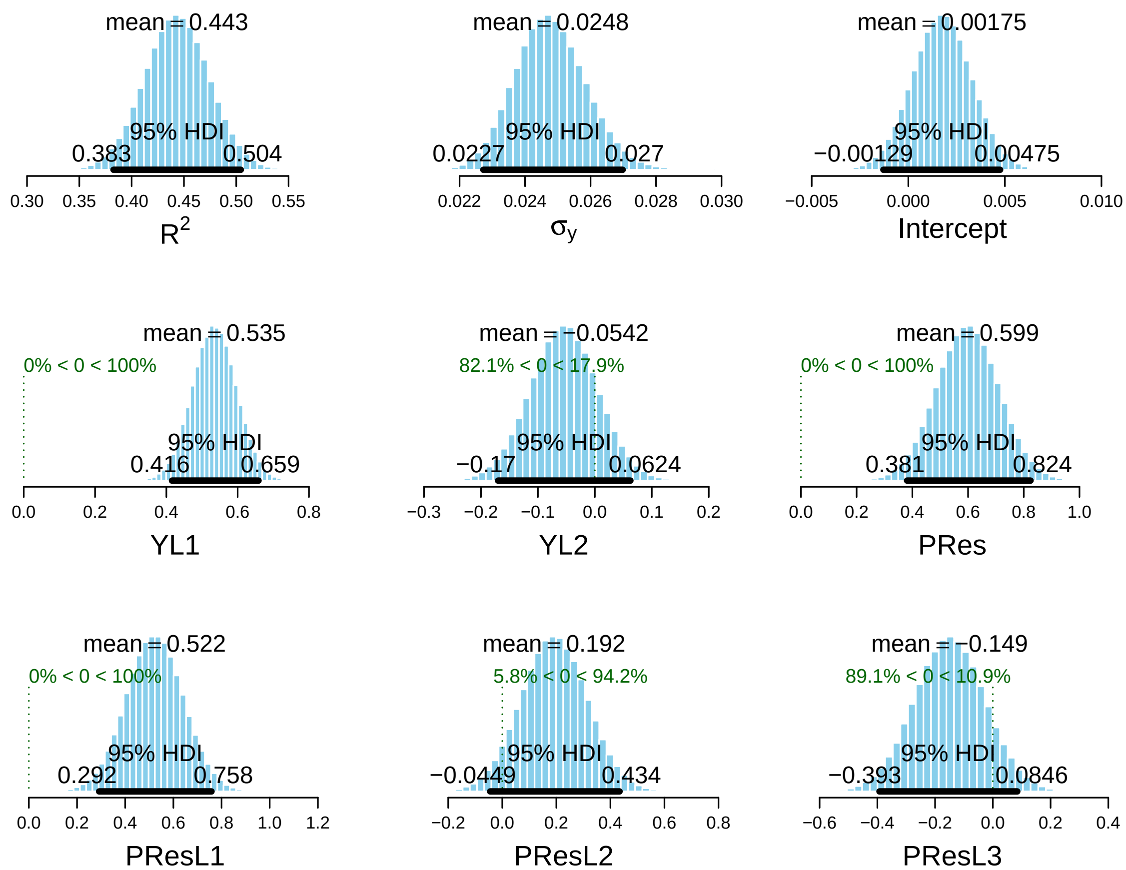

Figure 3: Posterior distributions of the parameters in Equation 2 using a normal likelihood.

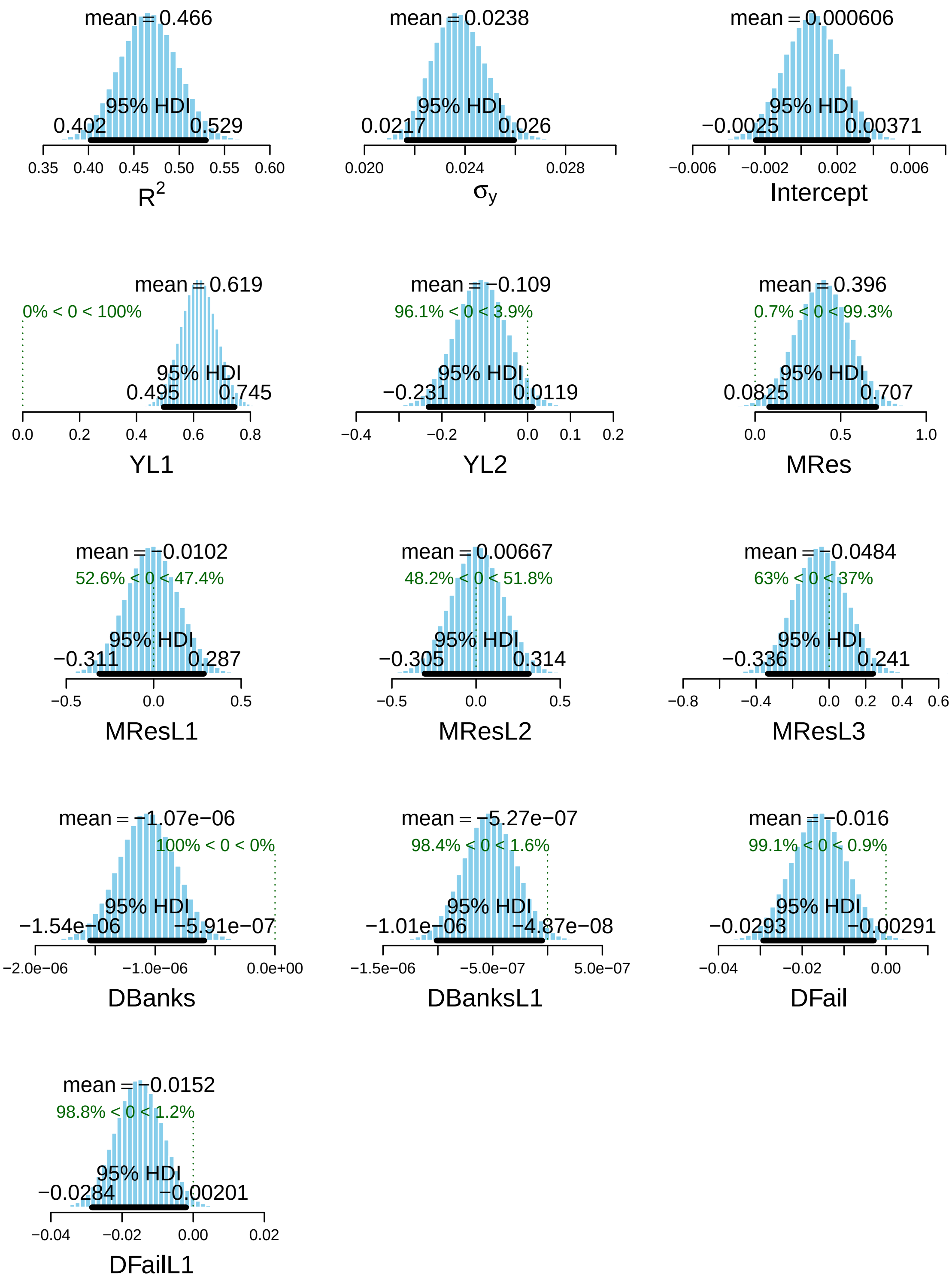

Figure 4: Posterior distributions of the parameters in Equation 3 using a normal likelihood.

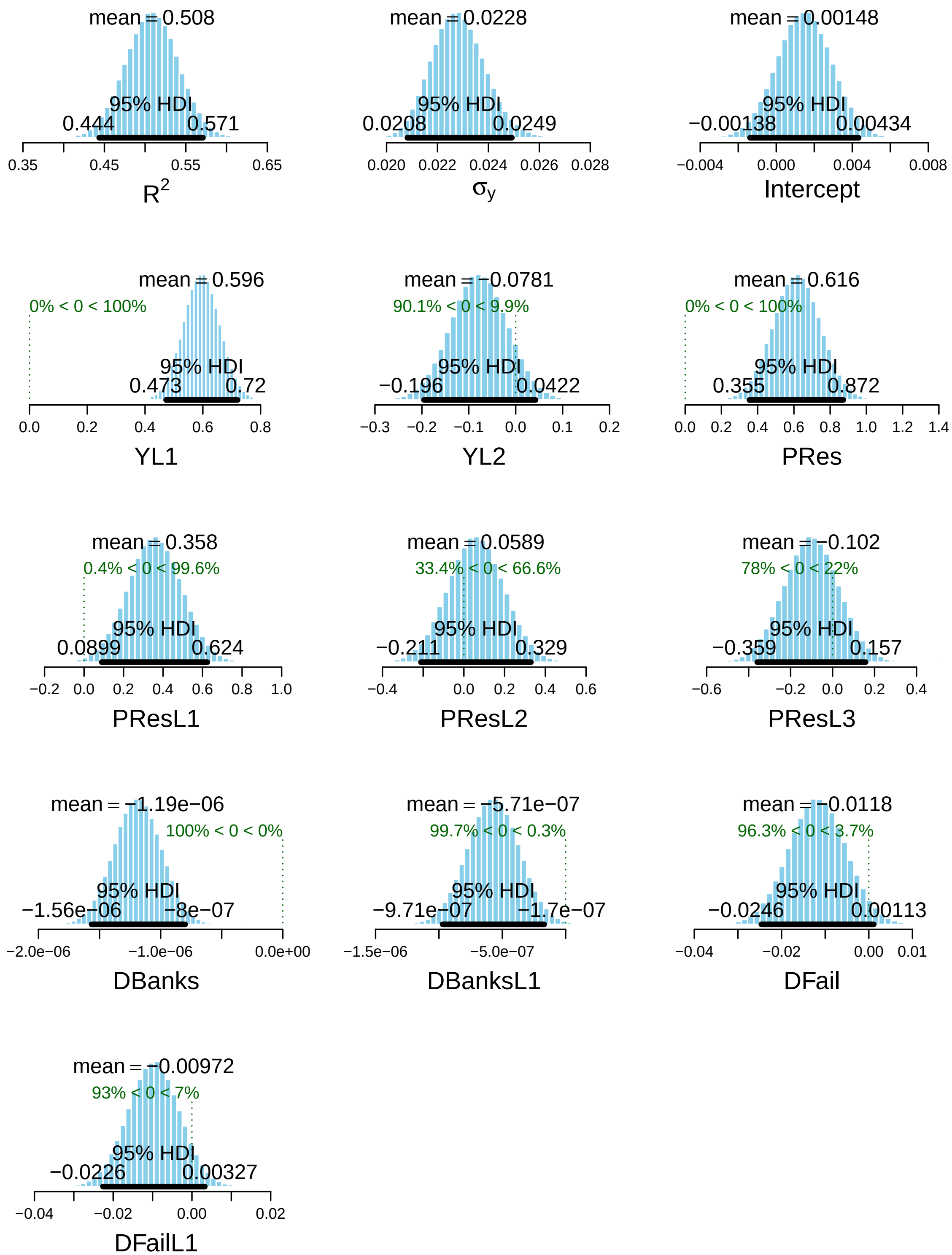

Figure 5: Posterior distributions of the parameters in Equation 4 using a normal likelihood.

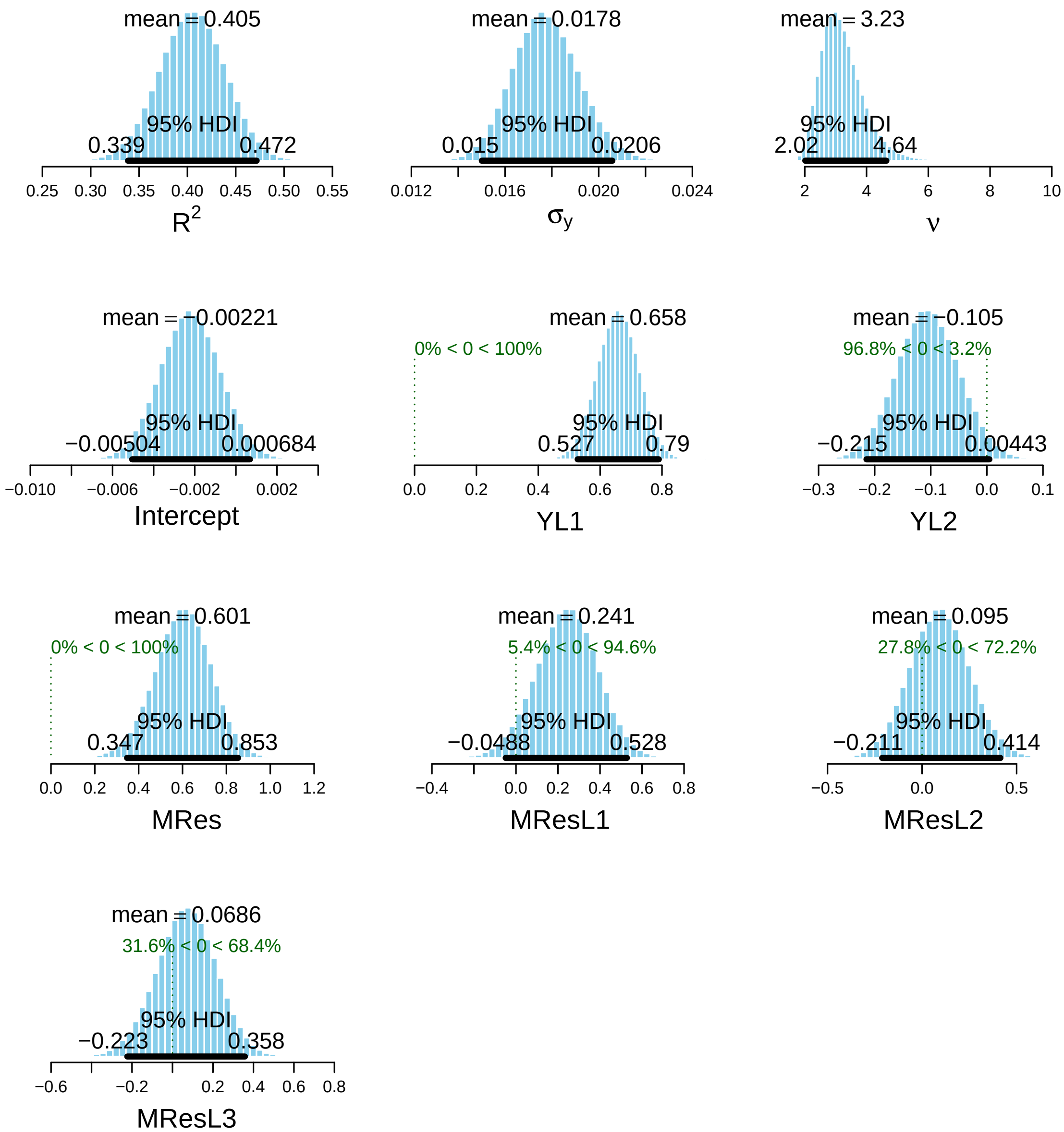

Figure 6: Posterior distributions of the parameters in Equation 1 using a t-likelihood.

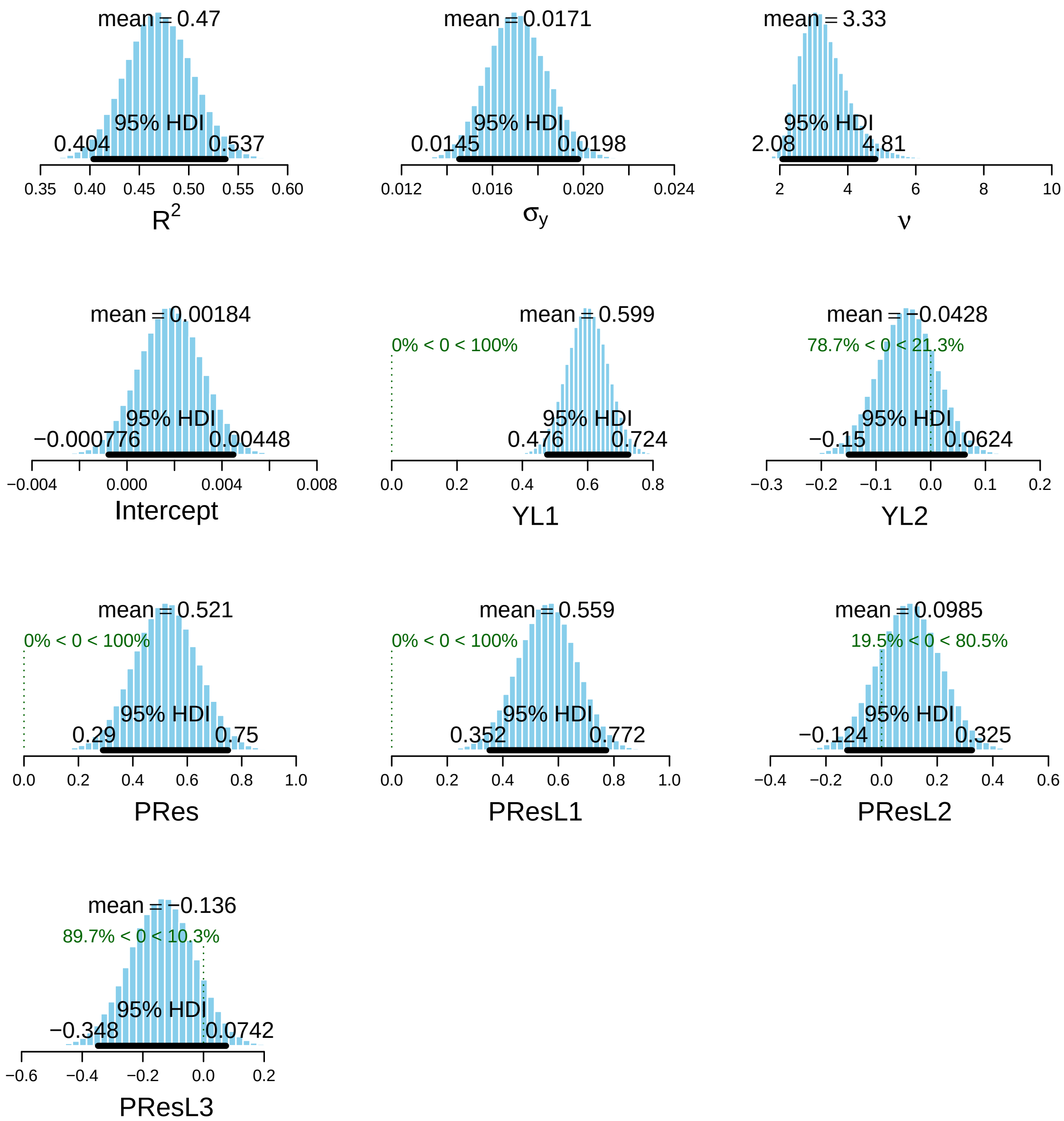

Figure 7: Posterior distributions of the parameters in Equation 2 using a t-likelihood.

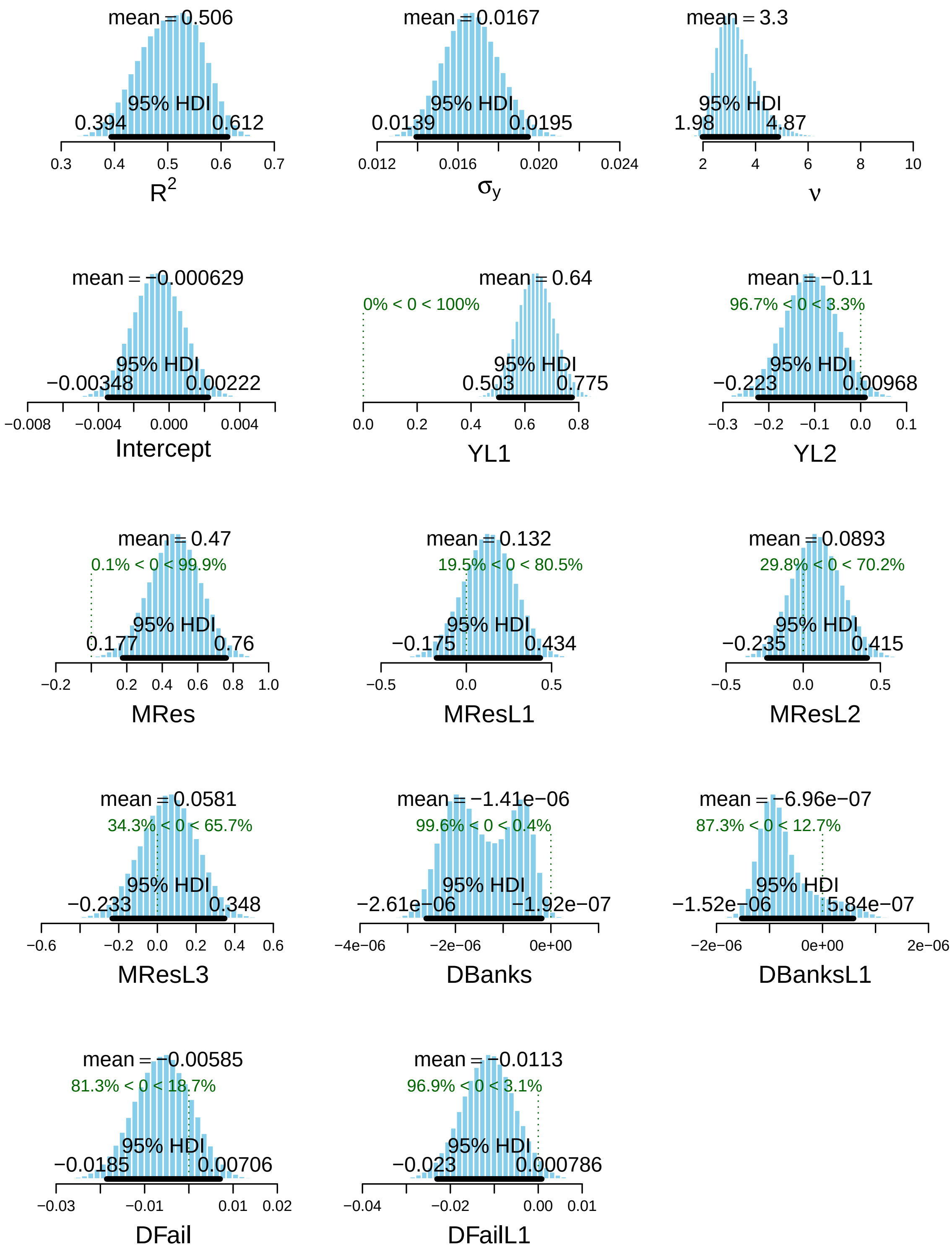

Figure 8: Posterior distributions of the parameters in Equation 3 using a t-likelihood.

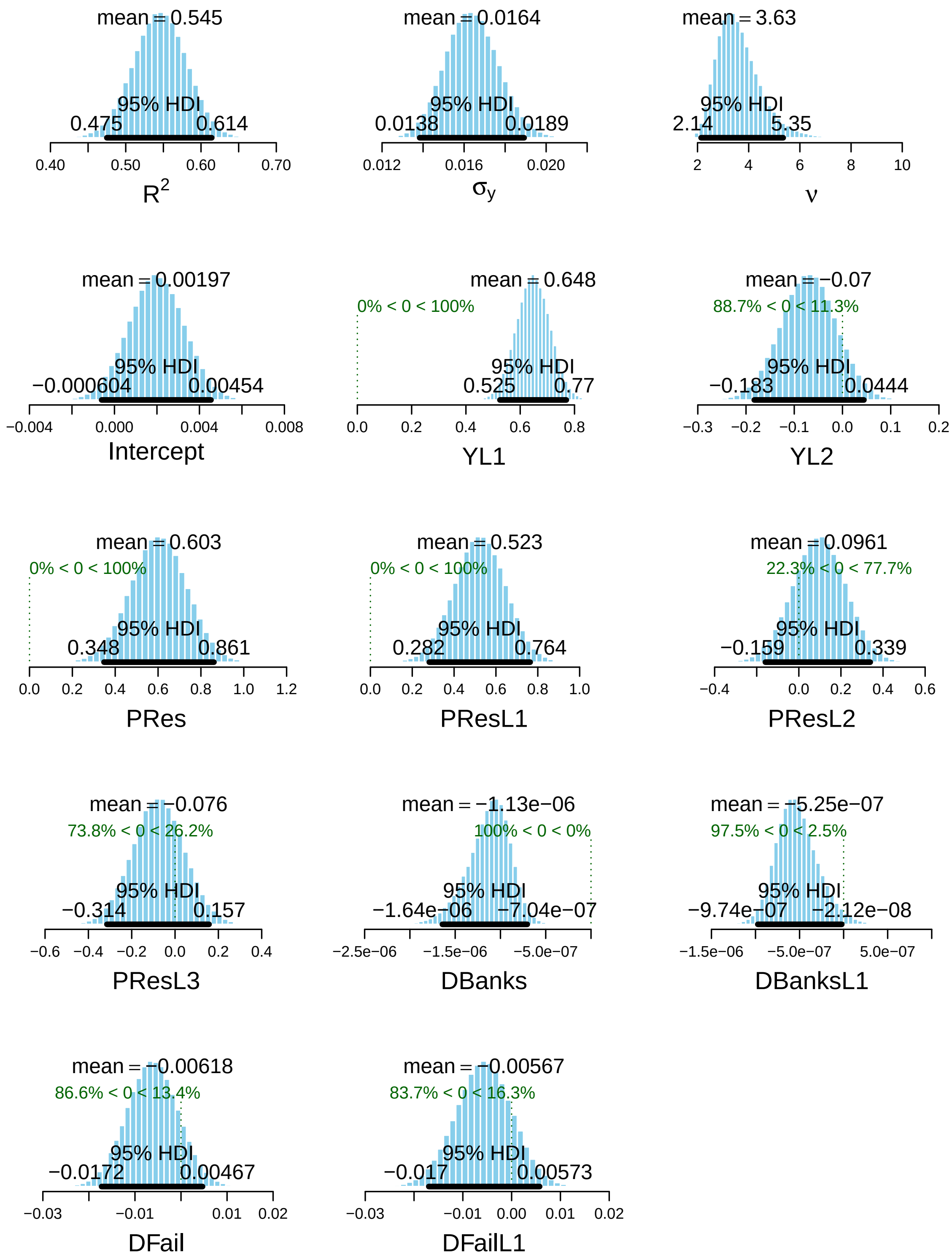

Figure 9: Posterior distributions of the parameters in Equation 4 using a t-likelihood.

References

[1] B. Bernanke, Non-Monetary Effects of the Financial Crisis in the Propagation of the Great Depression, *The American Economic Review*, **73**(3), (June, 1983), 257-276.

[2] J. Miron and N. Rigol, Bank Failures and Output During the Great Depression, NBER *Working Paper in Monetary Economics*, **19418**, (September, 2013), http://www.nber.org

[3] J. Mason, A. Ali , and J. Kolari, The Stock of Closed Bank Deposits, Duration of Credit Channel Effects, and the Persistence of the U.S. Great Depression, ASSA *Proceedings Cliometric Society Sessions*, Boston, (January, 2000).

[4] O. Nasraoui, A Brief Overview of Robust Statistics, website, (December, 2013),
http://webmining.spd.louisville.edu/Websites/tutorials/RobustStatistics

[5] J. Christmas and R. Everson, Robust Autoregression: Student-t Innovations Using Variational Bayes, *IEEE Transactions on Signal Processing*, **59**(1), (January, 2013), 48-57.

[6] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*, Chapman & Hall/CRC, 2004.

[7] J. Geweke, Bayesian Treatment of the Independent Student-t Linear Model, *Journal of Applied Econometrics*, **8**(1), (September, 1993), 19-40.

[8] J. Kruschke, *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*, Academic Press, 2011.

[9] J. Kruschke, H. Aguinis and H. Joo, The Time Has Come : Bayesian Methods for Data Analysis in the Organizational Sciences, *Organizational Research Methods*, **15**(4), (October, 2012), 722-752.

[10] J. Fox and H. Weisberg, *An R and Companion to Applied Regression*, SAGE Publications, Inc, 2002.

[11] M. Friedman and A. Schwartz, *A Monetary History of the United States, 1867-1960*, Princeton University Press, 1971.