MODELING AND CALIBRATING BANKS' DEMAND DEPOSITS VS. ASSETS

Elena Cristina Cânepă[1], Cristina Serbănescu[2], Alina Petrescu-Nită[3]

This paper considers a continuous time stochastic mathematical model to analyze financial processes. The goodness-of-fit using several normality tests is performed and compared for the cases when banks' deposits and assets evolve as Brownian motions, geometric Brownian motions, Ornstein-Uhlenbeck processes, and geometric Ornstein- Uhlenbeck processes. By using the equation of a regression line, the deposits' parameters are taken to be functions of the mean asset sizes. We obtain that the best fit is achieved when using the geometric Ornstein-Uhlenbeck process to model banks' deposits and assets.

Keywords: Brownian motion, (geometric) Ornstein-Uhlenbeck process, normality tests, calibration, regression, banks deposits, asset sizes.
MSC2010: 91G70, 91B70

1. **Introduction**

Stochastic models for default timing have developed into two different directions: the structural approach and the reduced form approach. The structural approach models the default as the first time when an underlying process, following a continuous path, falls below liability. For example, the models of Black and Scholes [3] or Merton [11] take the underlying process to be a geometric Brownian motion.

In this paper we assume that banks' deposits follow a continuous path and assets are constant. We extend the model proposed by Chen and Mazumdar in [9] where one bank is characterized by its asset size and its demand deposits that evolve as a Brownian motion with drift. In [9], the dependence of the deposits on the asset size is theoretically modeled as a monotone function, based on interpretations of the previous results on banks' risk-averseness, information-asymmetry and deposit-taking costs from [1], [2], [10], [12].

[1]University Politehnica of Bucharest, Faculty of Applied Sciences, R417-418 Department of Mathematical Methods and Models, 313 Splaiul Independentei, 060042 Bucharest, Romania; email (corresponding author): elenacristina2@gmail.com

[2]University Politehnica of Bucharest, Faculty of Applied Sciences, R417-418 Department of Mathematical Methods and Models, 313 Splaiul Independentei, 060042 Bucharest

[3]University Politehnica of Bucharest, Faculty of Applied Sciences, R417-418 Department of Mathematical Methods and Models, 313 Splaiul Independentei, 060042 Bucharest

We model the link between assets and deposits from 'scratch', by calibrating the deposits to the data and applying a regression between deposits' parameters and assets. We obtain a realistic model for deposits' parameters as functions of asset sizes, extending the results from [5].

Our findings can be used to model and simulate an economy with many banks characterized by assets and deposits, as in [7], [8] and [4]. Policy makers can use such models to test the impact of a particular policy on a simulated system, before its implementation in the real market. Such an artificial market can be created, for instance, starting from the empirical investigations on the banks' asset sizes from [6] and simulating deposits that evolve according to the formulas obtained in this paper.

Our contributions are the following. We analyze the relation between demand deposits and assets for 1221 U.S. banks that report non-zero deposits and assets for a period of 10 years. In the current American fractional reserve requirement regime, banks are required to keep 10% of their demand deposits in an account with the Federal Reserve Bank. We find that the ratio between banks' mean demand deposits and mean asset sizes oscillates around 0.1, for all the banks in our data set. The calibration of the banks' demand deposits to a (geometric) Brownian motion shows that most of the banks have positive deposit drifts, which increase as the asset size increases. However, 75 (and 96) banks out of 1221 present negative deposit drifts when deposits are calibrated to Brownian motions and geometric Brownian motions, respectively. We also calibrate deposits to Ornstein-Uhlenbeck and geometric Ornstein-Uhlenbeck processes. Normality tests show that the best fit is given by the geometric Ornstein-Uhlenbeck process. We propose four models for an economy with banks that are characterized by deposits and assets, by using the equations of the regression lines applied on the banks' estimated deposit parameters vs. mean asset sizes. Therefore, the models are closely linked to the real data.

The article is structured as follows. Section 2 presents the general model for the evolution of the demand deposit process and the dependence on the asset size. Section 3 describes the data and presents preliminary data mining results on the relation between the mean deposits and the mean asset sizes as a regression over all the banks. Section 4 presents the results of the goodness of fit tests when we compare four models: (geometric) Brownian motion and (geometric) Ornstein-Uhlenbeck process. Section 5 investigates the relation between the estimated deposit parameters and the corresponding asset sizes.

2. **The general model**

We consider an economy with N banks. The probability space is one which allows N independent standard Brownian motions $(B_t^i)_{t\geq 0}$, $i = 1, N$:

$$(\Pi_{i=1,N}\Omega^i, \Pi_{i=1,N}F^i, P).$$

Each bank i is characterized by a constant asset size A^i and by a diffusion process that models its demand deposits over time $(D_t^i)_{t\geq 0}$:

$$dD_t^i = \mu(A^i, (D_s^i)_{0\leq s\leq t})dt + \sigma(A^i, (D_s^i)_{0\leq s\leq t})dB_t^i.$$

3. **Data mining on asset sizes vs. demand deposits**

In this section we present the data and the direct connection between mean asset sizes and mean demand deposits. We obtain that the ratio between mean deposits and mean asset sizes oscillates close to 0.1 for all banks in our sample (10% being also the reserve requirement percent). We also represent the mean assets and the mean deposits for all banks on a logarithmic scale and provide the equation of the regression line.

Our data is obtained from WRDS (Wharton Research Data Services).

For each commercial bank in the United States we used $RCFD2170$ (total asset sizes) and $RCON2210$ (demand deposit amounts- net of withdrawals) on the last business day of each quarter between March 1991 and December 2000. The total number of banks is 1221. We included in our data set only the banks which had recorded non-zero total asset size and demand deposit amount, for all 40 quarters between 1991 and 2000.

Our first attempt in studying the connection between deposits and asset sizes is to calculate, for each of the $N = 1221$ banks, a ratio between the mean value of its deposits and the mean value of its assets, over the 40 quarters. We denote the mean asset sizes (over 40 quarters) of the N banks in our data set as constants $A^1, A^2, ..., A^N$. Similarly, we denote the mean demand deposits (over 40 quarters) of the N banks in our data set as constants $D^1, D^2, ..., D^N$.

In figure 1 we represent the 1221 banks on the x-scale and the corresponding ratios $\frac{D^i}{A^i}$ (i.e. deposits/ assets) on the y-scale. It shows that the ratio deposits/ asset sizes tends to be very close to 0.1, which is probably related to the reserve requirement percent $q = 10\%$. Figure 1 also presents a plot of the residuals, which appear randomly scattered around zero. Therefore the regression line gives a satisfactory fit to the data, allowing us to make the following approximation of the value of mean deposits as a function of mean asset sizes:

$$D = 0.1A.$$

Figure 2 represents the mean asset sizes $(A^i)_{i=1,N}$ versus the mean deposits $(D^i)_{i=1,N}$, on a log-log scale. The regression line give the following

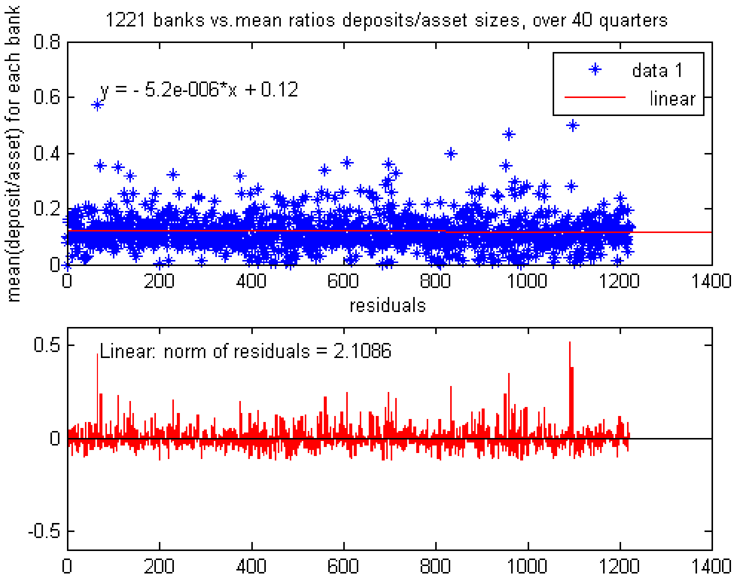

FIGURE 1. Mean ratios deposits/assets for all 1221 banks, March 1991 to December 2000

relation between mean assets A and mean deposits D:

$$D = \frac{A^{0.89}}{e}$$

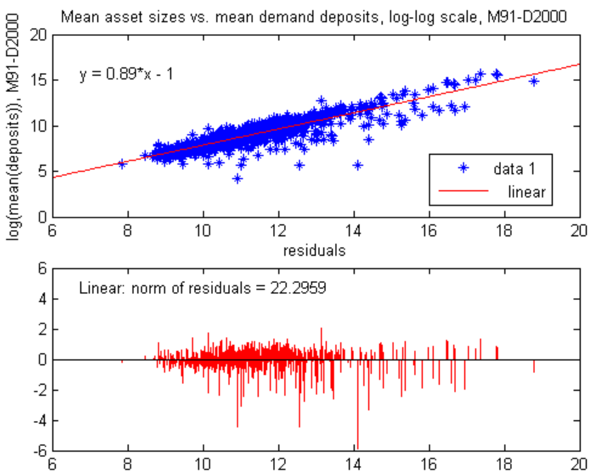

FIGURE 2. Mean asset sizes (A) vs. mean demand deposits (D), on a log-log scale

4. Modeling and calibrating demand deposits

We model and calibrate the demand deposit process as a Brownian motion with drift, a geometric Brownian motion, an Ornstein-Uhlenbeck process and a geometric Ornstein-Uhlenbeck process. The probability space is rich enough to allow N independent standard Brownian motions $(B^i_t)_{t\geq 0}$, $i = \overline{1, N}$, as in section 2. The question is which model fits the demand deposits best? First we calibrate the data to each model and then we test the goodness-of-fit.

4.1. Calibration

We present the formulas used for the calibration of the demand deposits in each model. For more insight on the discussed models we refer to [13]. Let $(D_t)_t$ be the demand deposit process for an arbitrary bank and $(B_t)_{t\geq 0}$ be a standard Brownian motion. Our data consists of $m = 40$ time-records of demand deposit amounts for each of the 1221 banks: $(D^i_t)_{i=\overline{1,N}; t\in[0=t_1<t_2<...<t_m=T]}$. We consider that demand deposits are recorded at m equidistant moments of time, i.e. $t_j - t_{j-1} = \delta$, for all $j = \overline{2, m}$. Consequently, $t_m - t_1 = (m-1)\delta$.

4.1.1. *Demand deposits as Brownian motions with drifts.* We consider that $(D_t)_{t\geq 0}$ follows a Brownian motion (BM) with drift μ and volatility $\sigma > 0$:

$$D_t = \mu t + \sigma B_t.$$

We consider the following estimates for the drift and for the volatility:
$\hat{\mu} = \frac{D_{t_m} - D_{t_1}}{t_m - t_1}$; $\hat{\sigma}^2 = \sum_{j=2}^{m} \frac{(D_{t_j} - D_{t_{j-1}})^2}{t_m - t_1} - \frac{1}{m-1}\frac{(D_{t_m} - D_{t_1})^2}{t_m - t_1}$.

One can check that $E[\hat{\mu}] = \mu$ and $Var[\hat{\mu}] = \frac{\sigma^2}{t_m - t_1}$. $E[\hat{\sigma}^2] = \sigma^2 - \frac{\sigma^2}{m-1}$.

4.1.2. *Demand deposits calibrated as geometric Brownian motions.* We consider that $(D_t)_t$ follows a geometric Brownian motion (GB) with drift μ and volatility $\sigma > 0$, i.e. $D_t = D_0 e^{\mu t + \sigma B_t}$, where $D_0 > 0$.

Then $(D_t)_{t>0}$ must satisfy the following stochastic equation (easy to prove using Ito's lemma):

$$dD_t = \mu D_t dt + \sigma D_t dB_t.$$

Since demand deposits are recorded at m equidistant moments of time, a unbiased drift estimator is: $\hat{\mu} = \frac{ln(D_{t_m}) - ln(D_{t_1})}{t_m - t_1}$.

A good volatility estimator is
$\hat{\sigma}^2 = \sum_{j=2}^{m} \frac{(ln(D_{t_j}) - ln(D_{t_{j-1}}))^2}{t_m - t_1} - \frac{1}{m-1}\frac{(ln(D_{t_m}) - ln(D_{t_1}))^2}{t_m - t_1}$.

Similarly as in the Brownian motion case, we have that: $Var(\hat{\mu}) = \frac{\sigma^2}{t_m - t_1}$ and $E[\hat{\sigma}^2] = \sigma^2 - \frac{\sigma^2}{m-1}$.

4.1.3. *Demand deposits calibrated as Ornstein- Uhlenbeck processes.* We assume that the demand deposit process $(D_t)_{t\geq 0}$ follows an Ornstein-Uhlenbeck process (OU), i.e.

$$dD_t = \theta(\mu - D_t)dt + \sigma dB_t,$$

where $\theta > 0$ is the mean reversion rate, μ is the mean, and $\sigma > 0$ is the volatility.

We obtain that $ED_t = D_0 e^{-\theta t} + \mu(1 - e^{-\theta t})$ and $Var(D_t) = \frac{\sigma^2}{2\theta}$.

The parameter estimators can be found using a least square regression:

$$D_{t_j} = D_{t_{j-1}} e^{-\theta\delta} + \mu(1 - e^{-\theta\delta}) + \sigma\sqrt{\frac{1 - e^{-2\theta\delta}}{2\theta}} N_{(0,1)} = aD_{t_{j-1}} + b + c\epsilon.$$

We denoted: $\epsilon_i = N_{(0,1)}$ i.i.d. standard normal random variables, $\delta = t_j - t_{j-1}$, for all $j = \overline{2, m}$; $a = e^{-\theta\delta}$; $b = \mu(1 - e^{-\theta\delta})$ and $c = \sigma\sqrt{\frac{1-e^{-2\theta\delta}}{2\theta}}$.

Then the model parameters are given by:

$\theta = -\frac{\ln a}{\delta}, \mu = \frac{b}{1-a}, \sigma = c\sqrt{\frac{-2\ln a}{\delta(1-a^2)}}$.

4.**1**.4. *Demand deposits calibrated as geometric Ornstein-Uhlenbeck processes.* The stochastic process $(D_t)_{t\geq 0}$ is a geometric Ornstein-Uhlenbeck process (GOU) if

$$d(\ln(D_t)) = \theta(\mu - \ln(D_t))dt + \sigma dB_t,$$

where $\theta > 0$ is the mean reversion rate, μ is the mean, and $\sigma > 0$ is the volatility.

The parameter estimators can be found similarly as in the OU case, using a least square regression: $\ln(D_{t_j}) = a\ln(D_{t_{j-1}}) + b + c\epsilon$.

ϵ's represent i.i.d standard normal random variables.

4.**2**. Goodness of fit tests

We compare the goodness of fit of four models for the demand deposit process (Brownian motion with drift, geometric Brownian motion, Ornstein-Uhlenbeck process, and geometric Ornstein-Uhlenbeck process). We use the Shapiro-Wilk test, known to be the most relevant for small samples. In [5] these models (with the exception of the geometric Ornstein-Uhlenbeck process) were tested on the same data using the Kolmogorov-Smirnoff test. However, as we present in section 4.**2**.1, the Shapiro test turns to be more appropriate for our data. We compare the results given by the Kolmogorov-Smirnoff test and the Shapiro test for all the four models. We obtain that the most adequate model among the proposed ones is the geometric Ornstein-Uhlenbeck process. Furthermore, our results show that there are banks for which deposits present negative drifts, signalizing a decreasing tendency in deposits.

We proceed by giving a quick presentation on the normality tests and on their power for our data set. The normality of a variable can be verified by the Shapiro-Wilk test (known as the best for small samples), the Kolmogorov-Smirnoff test (the most general), the Lilliefors test or the Anderson-Darling test [14]. It is relatively simple to apply these tests since they are implemented as functions in *matlab* (*kstest*, *lillietest*) and especially in *R*:

ks.test, *lillie.test*, *shapiro.test*, *ad.test*, *cvm.test*.

However, one should be cautious when applying these normality tests, it is possible that the failure to reject the null hypothesis to have been caused by an inadequate sample size. We need to compute, before proceeding to the data analysis, a power calculation that confirms that the sample size in that study was adequate for detecting a relevant difference.

4.**2**.1. *The power of the tests.* By definition, the power of the test is $1-\beta$, where β is the type II error (i.e. the probability of non-rejecting a false null hypothesis H_0). It is known that this error decreases as the sample size increases [14].

We perform the following experiments in order to find the necessary sample size. We generate samples from a non-Gaussian distribution and we apply normality tests to test the hypothesis that the simulated data follows a normal distribution. We estimate the type II error rate by repeating the experiment 10,000 times and computing a mean of the 10,000 p-values.

Nearly all the normality tests listed in the package *nortest* from R give a mean p-value $\leq .05$ for a sample size of $n = 20$, when simulating samples from a Cauchy distribution. Both the Anderson test and the Shapiro test do better than Lilliefors test, but the difference is small. However, when we use the normality tests to distinguish the difference between samples generated from a t-distribution $t(10)$, and the standard normal distribution, the power of the tests is very low. The Shapiro test does a lot better than the other methods implemented in the *nortest* package. Lilliefors test is the worst.

Therefore, when testing two hypotheses that are close to each other (in our case, BM/ GBM/ OU/ GOU) on a data that has 40 points (covering 10 years), the statistical evidence might be weak and inconclusive. However, if we get a larger data set, it will cover many regime changes. It would be risky to assume that the same model would hold for such a long period of time.

4.**2**.2. *Results of goodness-of-fit tests.* We have 1221 banks and 40 data points for each bank. Suppose that we have an i.i.d. sample $D_1, ..., D_{40}$ with some unknown distribution P and we would like to test the hypothesis that P is equal to a particular distribution P_0, i.e. decide between the following hypotheses: $H_0 : P = P_0$, $H_1 : P \neq P_0$.

Comparing the number of p-values that are $< .05$ when we apply the normality tests, we deduce that the best fit is given by the geometric Ornstein Uhlenbeck. The number of banks (out of 1221) for which the 39-data-points do not fit the Brownian motion, the Ornstein-Uhlenbeck (OU) process, the geometric Brownian motion, or geometric OU process are given in table 1.

Test	BM	OU	GBM	GOU
KS	132	93	72	38
Shapiro	702	523	523	324

Table 1. The number of banks (out of 1221 banks) for which the goodness-of-fit tests reject the null hypothesis

Because the Shapiro test is the most powerful, the numbers are a lot higher than those from the experiment using the Kolmogorov-Smirnov (KS) test. The result is clear: the GOU beats the other models and the BM is the worst fit.

5. Estimated deposit drifts and volatilities vs. asset sizes

In this section we study the connection between the parameters of the fitted models and the corresponding mean asset sizes. We provide the formulas for the regression line representing the estimated parameters vs. the mean asset size, across all the banks.

5.1. Estimated BM deposits parameters vs. asset sizes

Assuming the aggregate demand deposit of each bank follows a Brownian motion with drift (i.e. $dD_t = \mu dt + \sigma dB_t$), we obtain the estimated μ and σ according to the formulas presented in section 4.**1**.1. It turns out that there are 75 banks which have negative drifts and 1146 banks with strictly positive drifts.

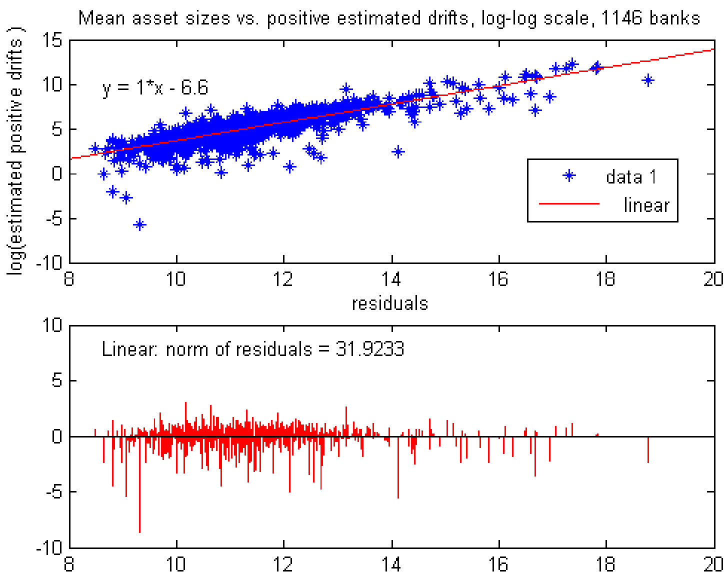

FIGURE 3. Mean asset sizes vs. estimated positive deposit drifts on a log-log scale (BM)

Figure 3 represents the mean asset sizes (A) vs. corresponding estimated drifts ($\hat{\mu}$) , on a log-log scale, for all the 1146 banks which have positive deposit drifts. The regression line gives an approximation for the positive drifts with respect to the asset sizes: $\hat{\mu} = Ae^{-6.6}$. The norm of the residuals (sum of the errors) is 31.

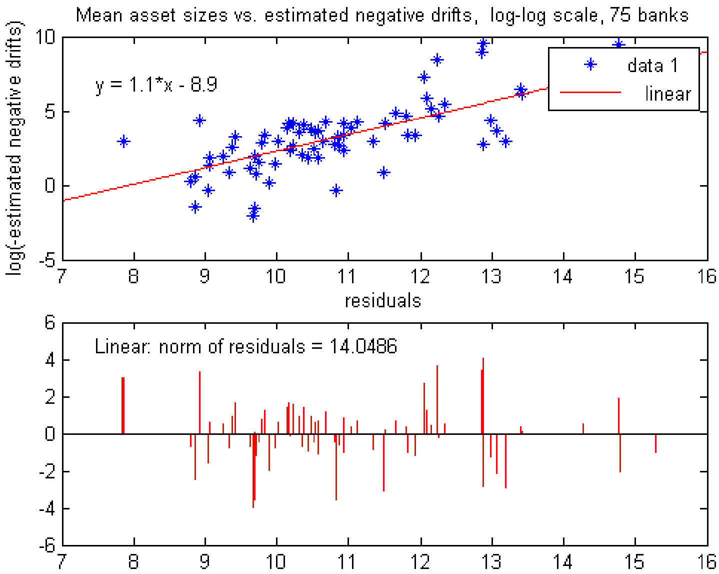

FIGURE 4. Mean asset sizes vs. estimated negative deposit drifts on a log-log scale (BM)

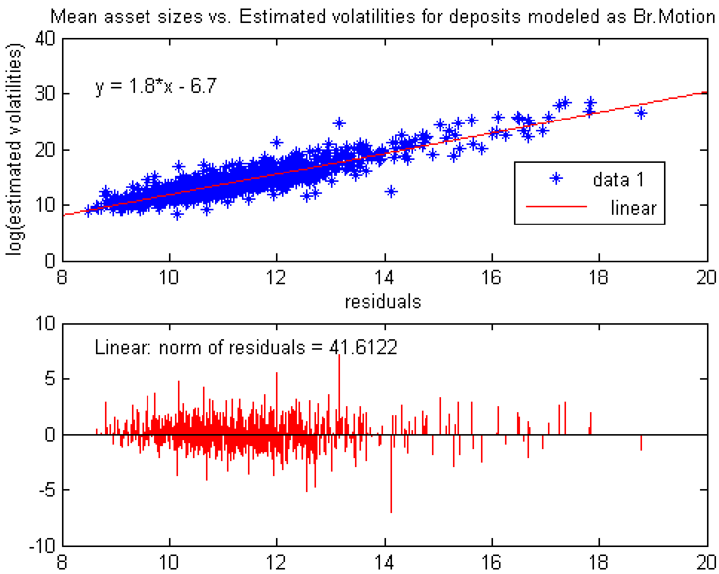

FIGURE 5. Mean asset sizes vs. estimated volatilities for the demand deposit processes (BM)

Figure 4 displays the mean asset sizes of the 75 banks which have negative deposit drifts vs. the opposites of the negative drifts, on a log-log scale. The coordinates of the regression line give a rough approximation of the negative drifts in terms of the mean asset sizes: $\hat{\mu} = -A^{1.1}e^{-8.9}$. The norm of the residuals is 14. Figure 5 represents the estimated volatilities ($\hat{\sigma}$) with respect to the mean asset sizes. The coordinates of the regression line give an approximation of the dependency of the volatilities on the mean asset sizes: $\hat{\sigma^2} = A^{1.8}e^{-6.7}$. The norm of the residuals is 41.

5.2. Estimated GBM/OU/GOU deposit parameters vs. assets

The procedure for the geometric Brownian motion case ($D_t = D_0 e^{\mu t + \sigma B_t}$) is similar to the BM case. In figure 6 the mean asset sizes A for all banks in economy are represented versus the estimated GBM demand deposit drifts ($\hat{\mu}$). The parameters are estimated according to the formulas from section 4.**1**.2. There are 96 banks which have negative drifts (out of the 1221 banks in our data set). The regression line gives an approximation for the drifts with respect to the natural logarithms of the mean asset sizes: $\hat{\mu} = 0.0033 \log(A) - 0.011$. The norm of the residuals is 1.2.

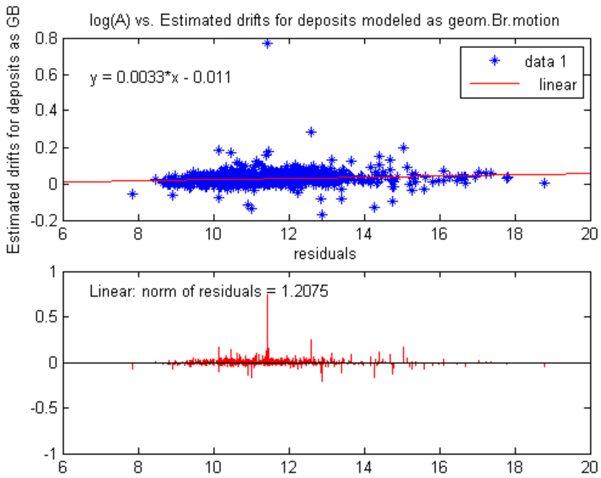

FIGURE 6. ln(mean asset sizes) vs. estimated GBM deposit drifts

In order to represent the asset sizes versus drifts on a log-log scale, we consider two disjoint sets of banks: the first set with the 1125 banks which have strictly positive GBM drifts and the second set with the 96 banks which have negative GBM drifts.

Table 2 displays the equations of the regression lines for the BM case and the GBM case. Table 3 displays the regression lines for the estimated deposit means and volatilities vs. asset sizes in the Ornstein-Uhlenbeck (OU) case and the geometric OU case (sections 4.**1**.3 and 4.**1**.4).

Figure 7 represents the mean asset sizes of the banks vs. corresponding mean reversion rates, on a log-log scale, in the GOU case.

Deposit model	*BM*	*GBM*
Pos. μ vs.A(residuals)	$\hat{\mu} = Ae^{-6.6}(31)$	
Neg. μ vs.A(residuals)	$\hat{\mu} = -A^{1.1}e^{-8.9}(14)$	$\hat{\mu} = -A^{0.29}e^{-8.1}(13)$
μ vs.A(residuals)		$\hat{\mu} = 0.0033\ln(A) - 0.011(1.2)$
σ vs.A(residuals)	$\hat{\sigma}^2 = A^{1.8}e^{-6.7}(41)$	$\hat{\sigma}^2 = A^{0.062}e^{-4.6}(33)$
No. neg. drifts	75	96

TABLE 2. The regression lines between the estimated parameters of deposits(fitted as BM, GBM) and the mean asset sizes

Deposit model	*OU*	*GOU*
Pos. μ vs. A (residuals)	$\hat{\mu} = A^{0.89}e^{-0.83}(32)$	$\hat{\mu} = A^{0.088}e^{1.2}(7)$
Neg. μ vs. A (residuals)	$\hat{\mu} = -A^{0.9}e^{-0.99}(18)$	$\hat{\mu} = -A^{0.29}e^{-8.1}(13)$
σ vs. A (residuals)	$\hat{\sigma}^2 = A^{0.91}e^{-3}(23)$	$\hat{\sigma}^2 = A^{0.022}e^{-2.1}(20)$
Number neg. drifts	89	6

TABLE 3. The regression lines between the estimated parameters of deposits fitted as OU/ GOU and the mean asset sizes

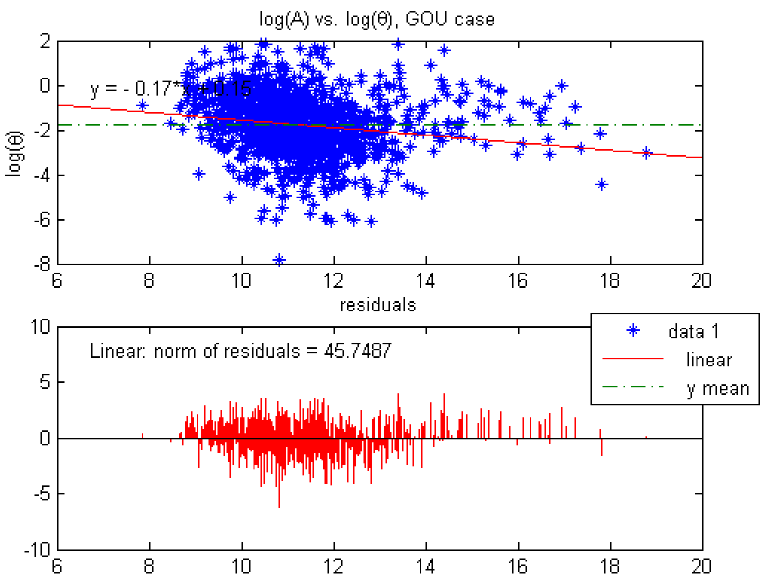

FIGURE 7. Mean asset sizes vs. estimated mean reversion rates for the demand deposit processes, log-log scale, GOU case

6. Conclusions

In this paper, we analyze the relation between banks' demand deposits and assets. We use quarterly data (covering 10 years) for 1221 U.S. banks. We obtain that the ratio between banks' mean demand deposits and mean asset sizes oscillates very close to 0.1, the reserve requirement percent. By calibrating banks' demand deposit to a Brownian motion, we find that most of the banks in our data set have positive deposit drifts. However, 75 and 96 banks out of 1221 are found to have negative deposit drifts when deposits are calibrated to Brownian motions and geometric Brownian motions, respectively. We provide the equations of the regression lines given by the banks' estimated deposit parameters vs. mean asset sizes. Such equations can be used for the simulation and the forecasting of a large banking system. We also calibrate deposits for Ornstein-Uhlenbeck (OU) and geometric OU processes. From the goodness-of-fit tests we obtain that the best fit for the deposits, among the proposed ones, is given by the geometric Ornstein-Uhlenbeck process. Future work might involve the adaptation of these models to biostatistics or biology.

REFERENCES

[1] *Allen L., Peristiani S., Saunders A.*, Bank Size, Collateral, and Net Purchase Behavior in the Federal Funds Market: Empirical Evidence,The Journal of Business,vol. 62, 1989

[2] *Allen, L. and Saunders, A.*, The large-small bank dichotomy in the federal funds market, Journal of Banking and Finance, vol. 10, no. 2, 219-230, 1986.

[3] *Black, F. and Scholes M.*,The pricing of options and corporate liabilities, Journal of Political Economy, vol.81,637-654,1973.

[4] *Canepa C., Pirvu T.*, A mathematical model and the optimal strategy in the transactions between one bank and the central bank,Proc. of the Romanian Academy(A),2019.

[5] *Canepa E.C. and Serbanescu C.*, Geometric Brownian Motion and Ornstein-Uhlenbeck process modeling banks' deposits, U.P.B. Sci. Bull., Series A, Vol. 72, Iss. 1, 2016.

[6] *Canepa C. and Canepa B.*, Dynamics of banks market concentration in U.S.A., Journal of Quantitative Methods, 11(4), 2016.

[7] *Canepa C. and Pirvu T.A.*, An Application of the Double Skorokhod Formula, Springer Procedings in Mathematics, Statistics, Vol.117, 2014.

[8] *Canepa Cristina*, Numerical Simulations of Defaults in Large Banking Systems, Ph.D. thesis, Carnegie Mellon University, 2012.

[9] *A. Chen and S. Mazumdar*, An instantaneous control model of bank reserves and Federal funds management, Journal of Banking and Finance, no. 16; 1992; 1073-1095.

[10] *Ho T. and Saunders A.*, A Micro Model of the Federal Funds Market, The Journal of Finance, vol.40, no.3, 977-988, 1985.

[11] *Merton, R.*, On the pricing of corporate debt: The risk structure of interest rates, Journal of Finance, vol.29, 449-470, 1974.

[12] *Rose and Kolari* A national survey study of bank services and prices arrayed by structures, the Journal of Bank Research, vol.16, 72-85,1985.

[13] *Shreve S.* Stochastic Calculus for Finance II, Springer, 2004.

[14] *Tuffery S.*, Data Mining and Statistics for Decision Making, John Wiley and Sons, 2011.