On the Stability of Money Demand

Robert E. Lucas Jr.
University of Chicago

Juan Pablo Nicolini
Federal Reserve Bank of Minneapolis and Universidad Di Tella.*

October 17, 2014

Abstract

To be done.

1. INTRODUCTION

The failure of Lehman Brothers in September, 2008 immediately led to a severe banking panic, a rush by banks to exchange privately issued cash substitutes for government issued or government guaranteed cash. The Federal Reserve responded to this situation by increasing the level of bank reserves from some \$40 billion on September 1 to \$800 b. by New Years Day. This single action was surely the main

*We thank Fernando Alvarez, Uberto Ennis, Finn Kydland, Ricardo Lagos, Francesco Lippi, Andy Neumeyer, Fabrizio Perry, Nancy Stokey, Pedro Teles and Randy Wright for helpful discussions and Manuel Macera and Serginio Sylvain for outstanding assistance.

factor in the resolution of the liquidity crisis by early 2009 and the ending of the decline in production after two quarters.

It is a remarkable feature of these events that none of the leading macroeconometric models—including the model in use by the Fed itself—had anything to contribute to the analysis of this liquidity crisis or of the Fed's response to it. None of these models had any role for bank reserves or for any other monetary aggregate or measure of liquidity. Central bankers, as always, used short interest rates as the only indicator of the stance of monetary policy but sometime in the 1990s they were joined by most influential monetary economists. A broad consensus was reached that no measure of "liquidity" in an economy was of any value in conducting monetary policy.

There were understandable reasons behind this consensus. Long standing empirical relations connecting monetary aggregates like M1, M2 and the monetary base to movements in prices and interest rates began to deteriorate in the 1980s and have not been restored since, as we document in Section 2. One objective in this paper is to offer a diagnosis of this empirical breakdown. A second is to propose a fix, to construct a new monetary aggregate that offers a unified treatment of monetary facts preceding and following 1980.[1]

To do this we need to get behind such broad aggregates as M1 and M2 and model the role currency and different kinds of deposits in the payment system in an explicit way. For this purpose we adapt the model of Freeman and Kydland (2000) (based on earlier work by Prescott (1987)) to consider the distinct roles of currency, reserves, and commercial bank deposits. This model proposes a banking "technology" that rationalizes the adding-up of different assets to form aggregates like M1, and does so with some realism. It treats currency and demand deposits as distinct assets

[1]Several economists have offerred useful diagnoses of the behavior of M1 and M2 since 1980 and proposed other monetary measures.These include Motley (1988), Poole (1991), Reynard (2004), Teles and Zhou (2005) and Ireland (2009).

and can readily be adapted to include other forms of liquidity. This is what we do in Section 3, where we introduce another monetary asset, money market deposit accounts (MMDAs), as a third important means of payment along side currency and demand deposits. The model rationalizes our use of a new, single monetary aggregate—we call it NewM1—that coincides with M1 prior to 1982 and includes MMDAs for the years since. In this section we treat banking activities as though they were conducted within each household, in order to situate banking activities within a general equilibrium framework. We decentralize the model in Section 4, obtaining a model of competitive banking firms.

In Section 5, we calibrate the model and compare its implications to U.S data for the free interest rates period (1915-1935 and 1983-2012). We show the model matches the behavior of the aggregate NewM1 remarkably well for the entire period. We also show that the model can reproduce some—but not all—of the features exhibited by the components of NewM1 in the data.

In the model described and applied in Sections 3-5, banking is treated as a free-entry, competitive industry, where banks are free to charge market interest rates on deposits. Of course, this description has never been literally accurate. In fact, among the provisions of the 1933 Glass-Steagall Act was the explicit prohibition—Regulation Q— of interest payments on bank deposits. In Section 6 we modify the basic model to provide an analysis of a banking system distorted by Regulation Q. We use this modified model to analyze the period with restricted interest rates (1936-1982) and discuss the important effects of Regulation Q on the U.S. financial system during the inflation of the 1970s and 80s. The model captures some - but not all - of the features of the Regulation Q period. Section 7 contains concluding remarks.

2. THE BEHAVIOR OF MONEY DEMAND

Figure 2.1 plots the ratio of M1 to GDP for the United States from 1915 to 1980 against the 3 month Treasury Bill rate. (Both series are annual averages) [2] This clear negative relation has been documented many times in empirical studies of money demand. These studies have typically ignored the distinction between the currency and deposit components of M1. In a sense, there was no need to address the composition issue since currency remained close to 25% of deposits during the period: See Figures 2.2 and 2.3. All three figures show the clear, negative relationship to the Tbill rate.

As Figure 2.4 shows, an important change occurred sometime in the early eighties. This figure extends Figure 2.1 from 1981 to 2012, with data from the latter period shown in red. Is this breakdown also reflected in the components of M1? The answer is shown in Figures 2.5 and 2.6: the breakdown in M1 is associated with a breakdown in the behavior of deposits, not in currency.

As described by Teles and Zhou (2005), the early eighties were a hectic period in terms of regulatory changes.[3] After the Great Depression, and following the Glass-Steagall Act and Regulation Q, commercial banks were prohibited from paying interest rates on bank demand deposits. Regulation Q was first relaxed slightly in 1980, when banks were allowed to pay limited interest on personal checking accounts (NOW accounts). In 1982 banks were allowed to issue interest-paying money market deposit accounts (MMDA), which could be held by some corporations as well as by households.

NOW accounts are essentially checking accounts in which certain restrictions, such as a minimum average balance, are imposed by banks as a condition of receiving interest on deposits. MMDA are less liquid, in the sense that a limited number of transactions (typically 6) are allowed per month. The Fed included NOW accounts

[2] The data base and its descrition is available upon request.

[3] See Teles and Zhou (2005) for details.

in M1, together with the traditional zero-interest checking accounts. However, it included MMDA with other savings accounts, which are part of M2.

A natural question, motivated by the coincidence in timing is: how did all these regulatory changes affect household decisions regarding the relative desirability of the different components of M1? Our candidate theory will be that the changes in the regulation in the early eighties substantially increased the availability of close substitutes for deposits, but not for cash. Once these close substitutes are allowed, our theory implies that they should be included in the monetary aggregate together with cash and traditional checking accounts. As a preview of our results, we plot, in Figure 2.7 the newly constructed aggregate—which we call NewM1—and its opportunity cost.[4]

Another major change, also emphasized by Teles and Zhou, occurred during the 90s, when the sweep technology was introduced. This essentially consists of software used by banks that automatically moves funds from checking accounts to MMDA's. This change substantially reduced the cost of moving funds between these accounts and made the MMDA more attractive relative to the demand deposits. Our theory keeps the technology parameters of banks fixed for the whole period considered, so it will be silent with respect to the effect of sweeps.[5] We will discuss them in more detail in Section 3.

3. THE MODEL

Freeman and Kydland considered a cash-in-advance model with two means of payment: currency and checks. We extend their model and incorporate money market

[4] The new aggregate contains assets that pay interest after 1980. The opportunity cost is defined as $\sum_j \alpha_j (i - i_j)$ where α_j is the share of asset j in the aggregate and $(i - i_j)$ is the difference between the TBill rate and the interest rate paid by asset j.

[5] A more detailed description of all the changes can be found in Teles ans Zhou (2005).

deposit accounts, the most important new addition to the stock of checkable deposits. To maintain the simplicity of the model, we will treat the traditional checking accounts and the NOW accounts as a single asset, just as Freeman and Kydland did. As in Figure 2.7 we will form the aggregate NewM1 by simply adding MMDAs to M1. But MMDAs and ordinary demand deposits coexist but are not perfect substitutes, so we need to model the distinct roles that these two assets play in the payment system. We do this by applying the Freeman-Kydland approach to three rather than two means of payment.

Their model was designed to introduce erratic money supply behavior into a real business cycle model. Our objective is to understand longer run or lower frequency relations. We therefore construct a deterministic stationary equilibrium with a constant technology for producing goods and a constant, perfectly foreseen growth rate in the supply of outside money. In this case the nominal interest rate r will be the Fisherian sum $\rho + \pi$ of the subjective, real discount rate and the money growth rate π. Since these components are exogenous in our set-up, we can express equilibrium relations in the steady state as functions of r. We view episodes of U.S. monetary history as temporary steady states that differ in a systematic way with differences in r. Then we compare the model's predictions to the actual behavior of prices and monetary aggregates determined by actual interest rate movements.

We turn to the details, beginning with preferences. Households consume a continuum of different goods in fixed proportions. Goods come in different "sizes", with production costs and prices that vary in proportion to size. Let $z > 0$ be the size of a good, and let $f(z)$ be a density on z, that indicates the number of goods of size z. Thus, $F(\widetilde{z}) = \int_0^{\widetilde{z}} f(z)dz$ is the fraction of goods of size less than or equal to $\widetilde{z}$. Let $\nu = \int_0^{\infty} zf(z)dz$. Consuming x means purchasing $\frac{z}{\nu}x$ units of each good of size z, so

$$x_t = \int_0^{\infty} x_t \frac{z}{\nu} f(z)dz = x_t \frac{\int_0^{\infty} zf(z)dz}{\nu}$$

All households have the common preferences over the perishable good x_t, of the form

$$\sum_{t=0}^{\infty} \beta^t U(x_t), \qquad \beta = \frac{1}{1+\rho}. \tag{1}$$

Each household has one unit of labor each period, to be divided between production and cash management. For now, we treat both activities as carried out by the household. A household must buy goods from other households and sell what it produces itself to others.

We next spell out the payments technology in which these three kinds of liquid assets are used. We assume, in the manner of Baumol (1952) and Tobin (1956), that households choose the number n of "trips to the bank" they take during a period. The availability of three payment methods complicates the nature of these trips. We follow Freeman and Kydland and assume that each period is divided into n stages that are identical in all respects. At the beginning of a year, a household begins with M dollars. These holdings are the economy's entire stock of base or outside money. These dollars are divided into currency C and bank deposits, $\theta^d D$ and $\theta^a A$, where D is the level of demand deposits and A is the level of MMDA's. These deposits are augmented with loans by the bank to increase the household's deposits, against which checks can be written, to levels D and A. During the first of the n subperiods, all of this currency and bank deposits will be spend on consumption goods. During this same initial period another member of the household produces and sells goods in exchange for cash and checks. At the end of the subperiod, producers visit a bank, checks are cleared, base money is divided as before, reserves, loans and deposits are renewed, and the situation at the beginning of the second subperiod replicates exactly the situation at the beginning of the first. This process is repeated exactly n times during the year.

We assume that all three payment modes are completely reliable, in the sense that sellers will accept payment in all of them, and that all modes involve the same "trip to the bank" time costs for buyers. Beyond these similarities, the three technologies have some different costs and benefits from a buyer's viewpoint. Money held as currency is subject to risk of theft or loss: We assume that a fraction $\tau \geq 0$ of each unit of currency held vanishes each period. Payments by either kind of deposit account are secure from these losses.[6] Both deposit types require a labor cost proportional to the number (not the value) of checks processed that will be assumed higher for the MMDA than for ordinary checking accounts. We denote these costs per check by k^d for demand deposits and k^a for MMDAs, and assume that $k^d < k^a$.

On the other hand, deposits require a fraction of cash being held idle as reserves. We assume that the fraction of cash required by demand deposits, θ^d, is higher than the one required by the MMDA's, θ^a. Thus, while payments with demand deposits are cheaper to process, they require a larger fraction of reserves and they may both be held in equilibrium. We will seek equilibria with two cutoff values $\delta > \gamma > 0$ such that sizes between 0 and γ are paid for by cash, sizes between γ and δ are paid for by demand deposits, and sizes larger than δ are paid for by MMDA's.[7]These cutoffs will be determined endogenously and will depend on the interest rate.

A household has one unit of time per period. The marginal product of labor—the single input—in units of the consumption good is given by the constant y. Total time allocated to both types of check processing is given by

$$x\left[k^d\left(F(\delta)-F(\gamma)\right)+k^a\left(1-F(\delta)\right)\right]$$

Bank trips add ϕn. The rest is allocated to production, so consumption will be given by

$$y(1-\phi n)-x\left[k^d\left(F(\delta)-F(\gamma)\right)+k^a\left(1-F(\delta)\right)\right]=x. \tag{2}$$

[6] Allowing for $\tau > 0$ is essential to have deposits held in equilibrium in our model of the regulation Q period. As we will see, it also implies that real money balances exhibit a satiation point at zero interest rates.

[7] We provide below a necessary and sufficient condition for both assets to be held.

The monetary base M will be assumed to grow at a constant rate π by means of lump sum transfers πM to each household. We let m denote a household's *relative* holdings of base money, so that in equilibrium $m = 1$. We renormalize money holdings every period. Let $p = P/M$ be the normalized price of goods and, similarly, let $c = C/M, d = D/M$ and $a = A/M$. For symmetry in the analysis, we define $\theta^c = (1-\tau)^{-1}$, so that θ^i, $i = c, d, a$ is the amount of base money one needs to hold for each dollar of payment made in mode i. Note that $\theta^c \geq 1$ and $\theta^d, \theta^a < 1$.

Required reserves are $\theta^d d + \theta^a a$ so base money is

$$c\theta^c + \theta^d d + \theta^a a \leq m. \tag{3}$$

We view each household as operating its own "bank" but still subject to a government-imposed reserve requirement. In the next section we will decentralize this allocation and explicitly assign different functions to households and banks. If we let

$$\Omega(\gamma) = \frac{1}{\nu} \int_0^\gamma z f(z) dz$$

be the fraction of total purchases paid for in cash, expressed as a function of the cutoff level γ, the cash constraints facing this consolidated household/bank are

$$nc \geq px\Omega(\gamma), \tag{4}$$

$$nd \geq px \left[\Omega(\delta) - \Omega(\gamma)\right], \tag{5}$$

$$na \geq px \left[1 - \Omega(\delta)\right]. \tag{6}$$

The law of motion for money balances is

$$m' = \frac{m + T + py(1-\phi n) - px\left(k^d\left(F(\delta) - F(\gamma)\right) + k^a\left(1 - F(\delta)\right) + 1\right) - (\theta^c - 1)c}{1+\pi}$$

Notice that the function F measures numbers of transactions while Ω measures numbers of dollars. Note also that the "lost" currency $\tau c \frac{1}{1-\tau} = \tau c\theta^c = (\theta^c - 1)c$ does appear as a negative item of the right. The variable T denotes the lump sum transfers, that include these cash balances lost and increases the total quantity of base money by $1+\pi$.

The household Bellman equation is

$$V(m) = \max_{x,c,d,a,n,\gamma,\delta} \{U(x) + \beta V(m')\}$$

subject to $(3)-(6)$. The first order plus envelope conditions evaluated at the steady state imply

$$\theta^c c + \theta^d d + \theta^a a = \frac{npy\phi - (\theta^c - 1)c}{r} \tag{7}$$

$$nk^d = \left((\theta^c - 1) + r\left(\theta^c - \theta^d\right)\right)\frac{1}{\nu}\gamma \tag{8}$$

$$\left(k^a - k^d\right)n = r\delta\frac{1}{\nu}\left(\theta^d - \theta^a\right) \tag{9}$$

The last two gives a relationship between the two thresholds, the interest rate and parameters of the model

$$\delta = \frac{\left(k^a - k^d\right)}{k^d}\frac{\left[\frac{(\theta^c-1)}{r} + \theta^c - \theta^d\right]}{\left(\theta^d - \theta^a\right)}\gamma. \tag{10}$$

A necessary and sufficient condition for the three assets to be held in equilibrium is that

$$h(r) \equiv \frac{\left(k^a - k^d\right)}{k^d}\frac{\left[\frac{(\theta^c-1)}{r} + \theta^c - \theta^d\right]}{\left(\theta^d - \theta^a\right)} > 1$$

Note that under our assumption that $\left(k^a - k^d\right) > 0$, MMDA's will be held when $\left(\theta^d - \theta^a\right) > 0$ and will disappear asymptotically when the term $\left(\theta^d - \theta^a\right)$ goes to zero from above.

In addition to these optimality conditions, a steady state equilibrium with positive interest rates must satisfy feasibility (2), the three cash-in-advance constraints $(4)-(6)$

with equality and the distribution of high power money condition (3) with equality and normalized high powered money equal to 1.

We can multiply each cash-in-advance condition by its corresponding reserve coefficient, add them up and use the high power money equilibrium condition to write

$$n = px\left[\theta^c\Omega(\gamma) + \theta^d\left[\Omega(\delta) - \Omega(\gamma)\right] + \theta^a\left[1 - \Omega(\delta)\right]\right]$$

Thus, the equations

$$(\theta^c - 1)\frac{px\Omega(\gamma)}{n} + r = npy\phi$$

$$k^d n = \gamma\frac{1}{\nu}\left[(\theta^c - 1) + r\left(\theta^c - \theta^d\right)\right]$$

$$py(1 - \phi n) = px\left[1 + k^d\left(F(h(r)\gamma) - F(\gamma)\right) + k^a\left(1 - F(h(r)\gamma)\right)\right]$$

$$n = px\left[\theta^c\Omega(\gamma) + \theta^d\left[\Omega(h(r)\gamma) - \Omega(\gamma)\right] + \theta^a\left[1 - \Omega(h(r)\gamma)\right]\right]$$

solve for p, x, n and γ. We can then use the last two to eliminate p and x, and obtain

$$\frac{n^2\phi}{(1 - \phi n)} = \frac{(\theta^c - 1)\Omega(\gamma) + r\left[\theta^c\Omega(\gamma) + \theta^d\left[\Omega(h(r)\gamma) - \Omega(\gamma)\right] + \theta^a\left[1 - \Omega(h(r)\gamma)\right]\right]}{\left[1 + k^d\left(F(h(r)\gamma) - F(\gamma)\right) + k^a\left(1 - F(h(r)\gamma)\right)\right]} \tag{11}$$

$$k^d n = \gamma\frac{1}{\nu}\left[(\theta^c - 1) + r\left(\theta^c - \theta^d\right)\right] \tag{12}$$

These two equations solve for the steady state values of (n, γ). Then, we can use (10) to solve for δ.

Given the solution, we can construct theoretical counterparts of observables, as follows: Money to output is given by

$$\begin{aligned} \frac{\theta^c c + d + a}{px} &= \frac{\theta^c px\Omega(\gamma) + px\left[\Omega(\delta) - \Omega(\gamma)\right]px\left[1 - \Omega(\delta)\right]}{pxn} \\ &= \frac{1 + (\theta^c - 1)\,\Omega(\gamma)}{n} \end{aligned} \tag{13}$$

while the ratios of currency to deposits and deposits to total deposits will be given by

$$\frac{c}{a+c+d} = 1-\Omega(\delta) \tag{14}$$

$$\frac{d}{a+d} = \frac{[\Omega(\delta)-\Omega(\gamma)]}{[1-\Omega(\gamma)]} \tag{15}$$

Analysis of the solution

There are only two relevant margins to be decided by households in this economy: the number of trips to the bank per unit of time, and the thresholds that determine the transactions to be made with cash and with each deposit type.

Equation (11) summarizes the optimal conditions for n, the number of trips to the bank. Note that if the thresholds γ, δ were exogenous, we could write the condition as

$$\frac{n^2}{1-\phi n} = rK$$

for some constant $K > 0$, which is the typical quadratic solution associated with a Baumol-Tobin type-model. It has only one positive solution $n(r)$ which increases as r increases, implying a negative relationship between real money balances—the inverse of n—and the nominal interest rate. Indeed, if the thresholds were fixed, the only relevant decision of households is the number of trips to the Bank and the model becomes a Baumol-Tobin model in which each of the transactional assets is proportional to the money supply.

On the other hand, equation (12) is the first order condition with respect to the threshold. Note that if the number of trips to the bank were exogenous, the solution for γ would be

$$\gamma\frac{1}{\nu}\frac{\left[(\theta^c-1)+r\left(\theta^c-\theta^d\right)\right]}{n} = k^d$$

The left-hand side is the proportional cost of using cash plus the foregone interest for a transaction of relative size $\frac{\gamma}{\nu}$. The right-hand side is the fixed labor cost of using demand deposits, which is independent of the size of the transaction. Thus, for transactions lower than γ, the direct plus foregone interests are lower than the fixed cost, so cash is the best asset to use.

When both n and γ are chosen by households, as we assume here, there are more possibilities. These are shown in Figure 3.1, where we plot the two functions: Equation (11) is the increasing quadratic one, with a positive intercept at $\gamma = 0$ and a finite asymptote as $\gamma \to \infty$.[8] Equation (10) is the straight line that goes through the origin. An equilibrium exists for $r \geq 0$.[9] Figure 3.2 illustrates the behavior of the equilibrium pair (n, γ) as r increases, shifting both curves upward. For the case shown in the figure, as the interest rate increases, n increases and γ decreases: Both real money balances and the ratio of currency to money balances decrease. But there is a possibility (not shown) that for some parameter values, either of the two effects of an increase in r could go the other way (though not both at the same time). In particular, the model allows for a non-monotonic relationship between real money balances and the short term interest rate. This could occur if the fraction of interest bearing deposits on money balances increases, possibly reducing the overall opportunity cost of money balances.

Notice that the interaction between the two decisions affects the value of the interest rate elasticity or real money balances. Consider again, in Figure 3.2, the movement from a "low" interest rate to a "high" interest rate. Imagine for a moment that the value of γ where fixed, at the solution for the "low" value of the interest rate. In this were the case, when the interest rate goes up to the "high" value, the solution for n would be given by the increasing and concave curve - the red curve - at that fixed

[8] The reason why the positive solution of the quadratic equation is increasing is because the right hand side of (12) is increasing in γ. And it asymptotes a finite value when $\gamma \to \infty$ since its limit is $\phi^{-1}\left(r\theta^c + (\theta^c - 1)\right)$.

[9] If $\theta^c = 1$, then $n \to 0$ as $r \to 0$, since equation (12) becomes the vertical axes, and real money balances do not have a satiation point.

value of γ - point A in Figure 3.2. But the value of γ will go down when the interest rate goes up. One could think of the effect of this adjustment in γ on the number of trips to the bank as the movement within the red curve, towards the intersection between the red curve and the red line - point B in Figure 3.2. This movement implies a weaker response of the number of trips to the Bank, given a change in the interest rate. i.e., a lower elasticity. The reason is that the portfolio recomposition (increasing the share of transactions made with deposits, that pay interest) implies that the overall opportunity cost of money balances goes up by less than what would be the case without the recomposition.

A note on technology growth.

An interesting feature of the solution of the system (11) and (12) is that is invariant with respect to changes in the marginal product of labor, y. Thus, the number of trips to the bank and the cutoff values will remain constant in a growth path, in which y grows overtime, as long as monetary policy is such that the interest rate, r , is constant. By (2), total consumption will grow at the same rate as productivity . This means that the number of transactions are also growing at the same rate of productivity, while the total number of hours devoted to do transactions is the same as before.

Therefore, in a balanced growth path, productivity is growing for both good producing activities and for transactions activities at the same rate. And the values of n, γ and δ are invariant to productivity. Thus, to the extend that productivity growth affects equally the two sector of the economy, the model can be used to study balanced growth paths like the one of the US during the last century.

4. DECENTRALIZATION

We have developed an equilibrium by treating production and banking services as different functions carried out within a single household. But in this constant returns environment, we can scale these activities up or down and decentralize them into banks and households. With free entry and competitive pricing—including interest rates on demand deposits—the decentralized equilibrium coincides with the results of Section 3. We then use the decentralized model to consider the effects of Regulation Q: the prohibition of interest payments on commercial bank deposits.

To carry this out, we need to allocate the two payment activities, "trips to the bank" ϕn and "check processing" $k(1-F(\gamma))$ between banks and households. As in Section 3 we assume that "trips" are a household activity that does not appear in the national accounts or as employment, and that banks process checks using labor hired from households at the equilibrium wage Py.[10] Households pay the bank a fee $Q^j = Mq^j,\ j = d, a$ per check processed from demand deposits and MMDAs respectively. The technology of check processing is unchanged.

As before, the normalized monetary base is divided into currency and reserves held against deposits. The deposit of θd entitles the depositor to withdraw d in checks. We treat this as a loan of $(1-\theta)\,d$ to the household by the bank, entitling the bank to $r\,(1-\theta)\,d$ at the end of the period. If the bank pays interest at a rate r_d on deposits the net interest cost to the household of setting up a deposit of size d is $(r\,(1-\theta)-r_d)\,d$. In a similar fashion, the net interest cost to the household of setting up a deposit of size a is $(r\,(1-\theta)-r_a)$. Finally, if the bank should make a profit this would be paid as a lump sum Π to households in their capacity as shareholders. The numbers $\left(q^d, r_d, q^a, r_a, \Pi\right)$ are taken as given by depositors. Their equilibrium values will be determined below. As before, monetary policy is executed via a lump sum transfer T.

The flow budget constraint facing the household is then

$$(1+\pi)\,m' = m + T + \left(r_d - r\left(1-\theta^d\right)\right)d + \left(r_a - r\left(1-\theta^a\right)\right)a - (\theta^c - 1)c$$

[10] These assumptions imply that all fixed cost of transactions acrrue within banks - it is inessential for the analysis that follows. We will allow for a more general sructure in the next sub-section.

$$+\Pi + py\left(1-\phi n\right) - px - q^d x\left(F(\delta)-F(\gamma)\right) - q^a x\left(1-F(\delta)\right)$$

The household's Bellman equation is

$$v(m) = \max_{x,c,d,a,\gamma,\delta,n}\left\{U(x)+\beta v(m')\right\}$$

subject to (3) and the cash constraints (4) – (6). As in Section 3 we obtain the first-order and envelope conditions, eliminate multipliers, and impose the steady state condition $m = m' = 1$ to obtain

$$n\frac{q^d}{p} = \frac{\gamma}{\upsilon}\left[(\theta^c-1)+r\theta^c-\left(r_d-r\left(1-\theta^d\right)\right)-r\theta^d\right]$$

$$n\frac{\left(q^d-q^a\right)}{p} = \frac{\delta}{\nu}\left[\left(r_d-r\left(1-\theta^d\right)\right)+r\theta^d-\left(r_a-r\left(1-\theta^a\right)\right)-r\theta^a\right]$$

$$\frac{n}{r}py\phi = r+\left[(\theta^c-1)c+d\left(r_d-r\left(1-\theta^d\right)\right)+a\left(r_a-r\left(1-\theta^a\right)\right)\right]$$

Bank profits will be

$$\begin{aligned}\Pi &= px\left[\Omega(\delta)-\Omega(\gamma)\right]\left(r(1-\theta^d)-r^d\right)+\left[F(\delta)-F(\gamma)\right]\left(\frac{q^d}{p}-k^d\right)\\ &\quad +px\left[1-\Omega(\delta)\right]\left(r(1-\theta^a)-r^a\right)+\left[1-F(\delta)\right]\left(\frac{q^a}{p}-k^a\right)\end{aligned}$$

If Banks set fees at the level of labor costs,

$$\frac{q^d}{p} = k^d, \quad \text{and } \frac{q^a}{p} = k^a,$$

and pay interest on deposits,

$$r_d = \left(1-\theta^d\right)r, \text{ and } \ r_a = \left(1-\theta^a\right)r$$

they will make zero profits. Then the first three conditions become

$$nk^d = \frac{\gamma}{\upsilon}\left[(\theta^c-1)+r\left(\theta^c-\theta^d\right)\right]$$

$$n\left(k^d-k^a\right) = \frac{\delta}{\nu}r(\theta^d-\theta^a)$$

$$\frac{n}{r}py\phi = r + (\theta^c - 1)c$$

which are the same as in Section 3, equations $(7) - (9)$. As the constraints are also the same, the equilibrium allocation is the same.

This equivalence between a decentralized equilibrium and a relatively centralized one is familiar, but of course it depends on prices being market-determined. In our application to the U.S. monetary system, though, this assumption seem's dubious for a good part of the period we analyze. As we mentioned, a notable feature of the 1933 Glass-Steagall Act was Regulation Q: the prohibition of interest payments on commercial bank deposits. This prohibition was in force from 1935 until 2011, though, as we argued, essentially bypassed by the deregulation of 1980-1982. The discussion above makes clear that restrictions on interest rates paid or on fees charged will have allocative effects. The precise nature of these effects will depend on the nature of Bank competition during the period.

What seems natural, though, given that the differential advantage of MMDA is the ability to pay higher interest rates than the other deposits, is that regulation Q would particularly affect the desirability of these deposits. This, we argue, explains why MMDAs were not developed within the banking sector during the regulation Q period. On the other hand, as they were allowed when interest rates were at their peak[11], they very quickly grew to more than 10% of GDP in less than two years, displacing the other deposits. This is our rationale for the use of NewM1 as the relevant aggregate for the whole period.

Note also that once NOW and MMDA accounts are allowed, the overall opportunity cost of money goes down, which should reduce the number of trips to the bank, and therefore increase real money balances. This implies that the regulation may affect the way a change in the interest rate affects real money balances. The case we analyze

[11] The high interest rates made regulation Q more costly which would explain why there was more pressure to change it.

in the next subsection explains this dependence in detail and shows that, indeed, the elasticity of real money balances with respect to the interest rate does depend on the regulation - a feature that is also present in the data.

5. CALIBRATION AND SIMULATION

The predictions of our version of the Freeman/Kydland model take the form given in equations (13)-(15), where the functions $n(r)$, $\gamma(r)$ and $\delta(r) = h(r)\gamma(r)$ are defined in (11), (12) and (10). For each interest rate level r, these three functions describe the corresponding steady-state behavior implied by the theory. In order to see how well this behavior matches up to U.S. time series we need to provide parametric descriptions of the functions F and Ω and then to provide values for the fixed parameters $\theta^c, \theta^d, \theta^a, k^d, k^a$ and ϕ. To do this, we draw on observations on currency, demand deposits, money market demand deposits, required reserve ratios, short term interest rate, the size of the banking sector, and nominal GDP.

For the distribution F of transaction sizes, we let the density be

$$f(z) = \frac{\eta}{(1+z)^{1+\eta}}, \quad \eta > 1$$

which implies that there are more small size transactions than large ones, as indicated by data on check payments.[12] This one-parameter density implies

$$F(\gamma) = 1 - \frac{1}{(1+\gamma)^{\eta}}$$

and

$$\Omega(\gamma) = \frac{\int_0^{\gamma} z f(z) dz}{\int_0^{\infty} z f(z) dz} = 1 - \frac{1+\gamma\eta}{(1+\gamma)^{\eta}}.$$

We add the parameter η to the list of constants to be calibrated. It will sometimes be convenient to make explicit the dependence of the functions F and Ω on the parameter η and write $F(\gamma, \eta)$ and $\Omega(\gamma, \eta)$.

[12] See, for instance, "The 2010 Federal Reserve Payments Study", pg. 27, http://www.frbservices.org/files/communications/pdf/research/2010_payments_study.pdf. .

For the values θ^c, θ^d and θ^a we assume $\theta^c = 1.01$, $\theta^d = 0.1$ and $\theta^a = 0.01$. The figure for θ^d is about the right magnitude for required reserves in the post-war period. We chose a much smaller (but still positive) reserve ratio for MMDAs. The value θ^c for currency would be 1 if currency were as safe as deposits. We chose a "loss rate" of 0.01 based on Alvarez and Lippi (2009).[13]

The three labor cost parameters k^d, k^a and ϕ and the parameter η of the transactions distribution are all assumed to be constants, but we have no direct evidence on any of them. We calibrate these indirectly, using the year 1984 as a benchmark. 1984 is the first year for which we have separate data on MMDA's, and two years after the MMDA's were permitted in the U.S. Two years was probably enough time for banks and others to adjust to the new instrument and to move funds from demand deposits to MMDA's.

The Tbill rate in 1984 was 9, 5% and the ratio of currency to total money holdings was 17%. Then for this year (14) implies

$$\frac{0.17}{1.01} = 1 - \Omega(\gamma\,(.095)\,,\eta). \tag{16}$$

In 1984 the values of demand deposits and MMDAs were approximately equal. This implies

$$1 - \Omega(\delta(.095), \eta) = \Omega(\delta(.095), \eta) - \Omega(\gamma(.095), \eta). \tag{17}$$

To complete the calibration, we estimated the fraction of resources used in carrying out transactions flows. Here we think of ϕn as household time, not included in measured GDP, and think of the labor costs per checks cleared, k^d and k^a, as bank employee time.[14] We can think of Phillipon's (2008) estimate of 0.05 as a bound on the fraction of GDP generated in the financial sector, but the fraction devoted to

[13] The number they find for Italy is 2%. We chose half of that number for the US and performed some robustness checks.

[14] Assuming that all the fixed costs are paid within Banks is not restrictive for the results of this section. This will not be the case in consdiering the regulation Q model in Section 6.

payment activities will be much less. We assume that total output spent on check processing at banks for both deposit types is 1% of GDP. In this case the following equation must hold for 1984:

$$k^d\left(F(\delta(0.095)) - F(\gamma(0.095))\right) + k^a\left(1 - F(\delta(0.095))\right) = 0.01 \tag{18}$$

We also assume that total time spent by households in trips to the bank is 1% of GDP.[15] Thus

$$\phi n(0.095) = 0.01 \tag{19}$$

Equations (16)−(19) plus the equilibrium conditions (10)−(12) evaluated at $r = 0.095$ then serve as 7 equations in the variables $\gamma(0.095), \delta(0.095), n(0.095)$ and the unknown parameters ϕ, k^d, k^a and η.

With the constant parameters determined, we then use the calibrated versions of (10) − (12) to solve for the functions $n(r), \gamma(r)$ and $\delta(r)$. Finally, we use the free scale parameter A as an intercept, chosen so as to give a reasonable fit to the empirical function

$$\frac{c+d+a}{px} = \frac{A}{n(r)}.$$

so that the theoretical curve crosses the point $(r = 0.06, m_1 = 0, 25)$. We chose to use annual data, and our cash-in-advance model is silent with respect to the choice of the time unit, and therefore silent with respect to the level of velocity. The free parameter A is therefore used to choose the average level of real money balances. Our theory is to be therefore seen as having implications only on the slope of the real money demand. It is for this reason that we are not using any information regarding the slope from the data and limit the calibration to use information of a single year, 1984.

The resulting values for the parameters are the following

[15] We will consider alternative targets and study the robusteness of the results to these numbers.

$$\phi = 0.0057 \quad \eta = 2.56 \quad k^d = 0.030 \quad k^a = 0.049$$

Figure 5.1, shows the evolution of the cash to deposit ratio predicted by the model when we fit in the observed interest rates from 1984 to 2012 in the black line with crosses, and the observed value for the cash to deposit ratio for the same period, in the blue solid line. This was a period with a downward trend in interest rates, but with very large high frequency fluctuations in the short run interest rate.

While it appears that the model does capture the overall increasing pattern of the cash to deposit ratio, the most noticeable feature is that the data exhibits much less volatility that the model. At this level, the model clearly fails. But in a sense, this is not completely surprising. It is well know that money demand models like the ones we consider in this paper fail to match the behavior of real money demand at high frequencies, in the sense that large, but temporary changes in the short run interest rate have relatively little impact on real money balances.[16]

To confirm that the model does partially capture the trend, Figure 5.2 reproduces the exercise, but compares the low frequency component of the cash to deposit ratio with the solution implied by the model when we feed into it the low frequency component of the short term interest rate. In both cases, we filtered the data using the Hodrick-Prescott filter with a smoothing parameter of 100. As it is clear from the graph, the model does capture remarkably well the upward trend exhibited in the data for the first two decades, but it misses by a large margin in the last one, when interest rates were very low.

[16] M1 to GDP did not respond at all to the short-lived reductions in interest rates in 71-72 and 75-76. Interestingly, NewM1 does respond (though by a small quantity) to the temporary drops in 86-87 and 92-93. Our model, by focusing on steady states only, does not provide any insigth on this feature. Attempt at explaining this feature go back at least as early as the models of Rotemberg (1984) and Grossman and Weiss (1983).

Figure 5.3 is similar to 5.1, but it measures the ratio of demand deposits to money market demand accounts. As before, the model clearly misses at the high frequency level. The same disclaimer applies. As figure 5.4 shows, however, the model reproduces the trend in the data, but only till the middle of the nineties. From there on, the ratio of demand deposit to total deposits falls dramatically, while the theory predicts that it should keep on increasing. As we discussed in Section 2, we believe that the Sweep technology introduced in 1994 and that became very popular by the end of the 90's explains the significant decline of demand deposits relative to MMDA observed in the data. Our constant technologies model is unable to capture this feature.

In summary, when considering the behavior of the ratio of the components of NewM1, the model produces mixed results relative to the data.

Finally, in Figure 5.5, we plot the calibrated theoretical curve for real money balances as a function of the nominal interest rate, together with the data for the unregulated periods (1915-1935 and 1983-2008).[17] Recall that we chose the free parameter A so as the curve must go through the point $(0.06, 0.25)$ so the curve goes through the data by construction. However, the calibration only used data for a single year, so no information regarding slopes has been feed into the model. In spite of that, the curvature of the theoretical curve reproduces the data notably well.

We find remarkable that in spite of the mixed results regarding the ratio of the components of NewM1 after the middle of the 90's, no sign of instability is detected for aggregate money balances. Thus, the very large substitution between demand deposits and money market demand accounts the data exhibits since 1995, did not affect the relationship between the short term interest rate and money balances over

[17]Note that the theoretical model takes into account the fact that the choices of each asset are endogenous. Still, the relationship derived relates the monetary aggregate to the nominal interest rate, not to the opportunity cost of holding money. It clearly does take into account the substitutiotns within the aggregate implied by the model.

GDP since 1995 and 2012.

Interestingly enough, this is also the case in the model, which exhibits a strong robustness of the overall money balances to GDP to many parameters that affect very critically the behavior of each of the components, like the parameters of the banking technologies. We performed several robustness exercises, one of them is particularly suited for this discussion, so we presented it here. We simulated the model for 4 different values of the parameter k^a, while keeping constant all other parameters. Calibration 1 will refer to the one we used so far. Then, calibrations 2, 3 and 4 use the same parameters, except that k^a, that was 0.049 in the original calibration, will take values $0.0466, 0.0433$ and 0.0400 correspondingly. These are important changes, since the ratio of k^a/k^d goes from almost 5/3 to 4/3.

Figure 5.6 shows, for the low frequency only, an exercise as the one in Figure 5.4 . Calibration 1 is the original one, which tracks the data well for the first part of the sample. As we reduce the fixed cost of doing transaction with the money market demand accounts, relative to the demand deposits, the cutoff for transactions with MMDA goes down, and the ratio of demand deposits to total deposit goes down.[18] Figure 5.7, shows the effect of this parameter change on total balances. Intuitively, as the cost of making transactions with MMDA's goes down, the overall opportunity cost of money balances goes down and real money balances do go up. But the change is not nearly as dramatic as it is for the ratio of demand deposit to total deposits. It is hard to argue that data on the aggregate NewM1, can fruitfully be used to learn with precision the technological parameters of the transactional banking sector.

We also explored the robustness of the model to changes in the targets for the calibration, in particular, the fraction of GDP used in transactions and the value of time spent by households in transactions when interest rates are close to 10%.[19] We

[18]Note, on the other hand, that the behavior at very low rates does change very litttle. This implies that the model will have a hard time matching this ratio for near zero interest rates.

[19]The matlab program "robustness.m" on the website of the paper can be used to solve the model

used in our benchmark calibration a value of 1% in both cases. Very similar results are obtained for the equilibrium values of the model if both numbers are set to 2%, or if the fraction of GDP used in transactions is left at 1% and the value of time spent by households is varied from 1.25% to 0.75%. Some plots get better, others get worse, but the general message is the same: While the ability of the model to match the ratios between the components of NewM1 is mixed, the behavior of the aggregate is remarkably close to the data.[20]

In Figure 5.5, we plotted the theoretical curve, together with the data for the periods where regulation Q was not in place, so we did not include the data between 1935 and 1981. The reason is that the theory assumes that interest rates are not regulated. We now consider, in the next section, a possible model of bank competition with legal caps on the interest rates banks can pay. However, before doing that, it is useful to compare the data during the regulation Q period, with the theory developed so far, which we do in Figure 5.8. Clearly, the theoretical curve overpredicts money balances - alas, because of regulation, the theory is not the correct one. In other words, the elasticity of real money balances to the interest rates was higher during the regulation Q period than the one implied by the model (and observe during the unregulated period). As we will see in the next Section, this is consistent with the model.

---

for different values of the targets.

[20] In this model, the welfare cost of inflation as a function of lifetime-consumption, is given by the sum of those two shares. This implies that the model as it stands is not very usefull to think about that issue. This is related to the need of the free parameter A to match average balances, which means ww cannot use the mdel to separately identify the value for the number of trips to the bank per period, n. We only identify how this value changes when the nominal interest rate changes.

6. REGULATION Q

In this section, we propose a possible model of banking competition under regulation Q. As we argued at the end of Section 3, when MMDA are not allowed to pay a higher interest rate than other deposits, they lose their relative advantage, so we will ignore them in this discussion. We therefore consider only the deposits with the lowest fixed transaction cost. Regulation Q is interpreted as imposing a ceiling, $\bar{r} = 0$ on the rate banks can pay on deposits.

The key idea we pursue in this section is that if banks are unable to pay interest on deposits, they will compete by offering reduced check processing fees. We treat each transaction size as a different product, and the fees will be determined by the zero profit condition product by product. We however impose the restriction that fees cannot be negative. Thus, for a given transaction size, the higher is the interest rate, the lower will be the fee charged by the bank for that transaction, to the point in which the interest income suffices to fully cover the fixed transaction cost. For higher interest rates the Banks will make positive profits, so some form of entry restriction must prevail for this outcome to be consistent with an equilibrium. As suggested by this discussion, the way the fixed costs k^d are borne by the household or the bank - an inessential distinction for the model described in Section 3 - becomes important. Thus, we let $k^d = k^b + k^h$, where $k^b \geq 0, k^h \geq 0$.

The details are as follows. Suppose there are $xf(z)$ checks to be processed at transaction size z and $px\frac{z}{\nu}f(z)$ dollars spent. These flows require $pxf(z)\frac{z}{\nu}\frac{1}{n}$ in deposits and entail $pxf(z)k^b$ in check processing costs. When the excess of interest income over cost at the transaction size z is positive, the depositor is charged no fee and the bank collects

$$\pi(z) = pxf(z)\left[r(1-\theta^d)\frac{1}{\nu}\frac{z}{n} - k^b\right]$$

in profit income. When it is negative, the depositor is charged with processing fees equal to

$$pxf(z)\left[k^b - r(1-\theta^d)\frac{1}{\nu}\frac{z}{n}\right]$$

for all checks of size z issued, and the bank collects no profit income. Thus, for a check of size z, the per-check fee in equilibrium must satisfy

$$\frac{q(z)}{p} = \max\{\left[k^b - r(1-\theta^d)\frac{1}{\nu}\frac{z}{n}\right], 0\} \tag{20}$$

Consumers see contracts such that

$$r(1-\theta^d)\frac{px}{n}\left[1-\Omega(\gamma)\right]$$

is the amount of money that cost them to borrow deposits equal to $\frac{px}{n}\left[1-\Omega(\gamma)\right]$. In addition, they take as given the function $q(z)$ that indicates the fee they pay for each check of size z. The total transactions cost for them is given by

$$\int_{\gamma}^{\infty}\left(\frac{q(z)}{p}+k^h\right)pxf(z)dz$$

Adding the two components, we define a function $M(n,\gamma)$ by

$$M(n,\gamma) = \int_{\gamma}^{\infty}\left(\frac{q(z)}{p}+k^h\right)f(z)dz + r(1-\theta^d)\frac{1}{n}\left[1-\Omega(\gamma)\right] \tag{21}$$

be the total cost to consumers of using deposits.

A depositor's Bellman equation in this situation is

$$V(m) = \max_{x,\gamma,n}\{U(x)+\beta V(m')\}$$

subject to the cash constraint

$$px\left(\theta^c\Omega(\gamma)+\theta^d\left[1-\Omega(\gamma)\right]\right) \leq mn,$$

and where

$$m' = \frac{m+\pi+\Pi+py\left(1-\phi n\right)-px-\frac{(\theta^c-1)px\Omega(\gamma)}{n}-pxM(n,\gamma)}{1+\pi},$$

where Π represents profits banks may make in equilibrium. As before, the first order and envelope conditions evaluated at the steady state, plus all other equilibrium conditions can be arranged to obtain two equations that solve for n and γ,

Using the last two to eliminate p and x, we obtain

$$n\left(\frac{q(z)}{p}+k^h\right)=\gamma\frac{1}{\nu}\left(\theta^c-1\right)\left(1+r\right)$$

which, using the value for the fee, (20) can be written

$$n\max\{\left[k^b-r(1-\theta)\frac{1}{\nu}\frac{z}{n}\right]+k^h,k^h\}=\gamma\frac{1}{\nu}\left(\theta^c-1\right)\left(1+r\right)$$

and

$$\frac{n^2\phi}{(1-\phi n)}=\left[\frac{r\left[\theta^c\Omega(\gamma)+\left[1-\Omega(\gamma)\right]\right]+(\theta^c-1)\Omega(\gamma)}{\left[1+k(1-F(\gamma))\right]}\right]$$

Analysis of the solution.

For the first equation, assume $\left[k^b-r(1-\theta)\frac{1}{\nu}\frac{\gamma}{n}\right]>0$, which means that the equilibrium fee is always positive at the cutoff value γ, $\frac{q(\gamma)}{p}>0$. Then, the first equation becomes

$$n\left[k^b-r(1-\theta^d)\frac{1}{\nu}\frac{\gamma}{n}+k^h\right]=\gamma\frac{1}{\nu}\left(\theta^c-1\right)\left(1+r\right)$$

or

$$n\left(k^b+k^h\right)=\gamma\frac{1}{\nu}\left[r\left(\theta^c-\theta^d\right)+\left(\theta^c-1\right)\right]$$

which is the same equation we obtained in Section 3 (see equation (12)).

On the other hand, if $\left[k^b-r(1-\theta^d)\frac{1}{\nu}\frac{\gamma}{n}\right]\leq 0$, we obtain

$$nk^h=\gamma\frac{1}{\nu}\left[r\left(\theta^c-1\right)+\left(\theta^c-1\right)\right]$$

The question is now, when will $\left[k^b-r(1-\theta^d)\frac{1}{\nu}\frac{\gamma}{n}\right]>0$?

Note that in this case,

$$n\max\{\left[k^b-r(1-\theta^d)\frac{1}{\nu}\frac{\gamma}{n}\right]+k^h,k^h\}=\gamma\frac{1}{\nu}\left(\theta^c-1\right)\left(1+r\right)$$

becomes

$$\frac{\left(k^b+k^h\right)}{\left[\frac{1}{\nu}\left(\theta^c-1\right)\left(1+r\right)+r(1-\theta^d)\frac{1}{\nu}\right]}=\frac{\gamma}{n}$$

Replacing the ratio $\frac{\gamma}{n}$ above, we obtain

$$\frac{\left(\theta^c-1\right)}{\left(1-\theta^d\right)}\frac{k^b}{k^h}>\frac{r}{(1+r)}.$$

Thus, for low enough values of the interest rate, such that

$$\frac{r}{(1+r)}<\frac{\left(\theta^c-1\right)}{\left(1-\theta^d\right)}\frac{k^b}{k^h}, \tag{22}$$

the equilibrium value of the fee at the cutoff value is positive, $\frac{q(\gamma)}{p}>0$, and the solution of the system is given by

$$n\left(k^b+k^h\right)=\gamma\frac{1}{\nu}\left[r\left(\theta^c-\theta^d\right)+\left(\theta^c-1\right)\right] \tag{23}$$

$$\frac{n^2\phi}{(1-\phi n)}=\left[\frac{r\left[\theta^c\Omega(\gamma)+\left[1-\Omega(\gamma)\right]\right]+(\theta^c-1)\Omega(\gamma)}{\left[1+k(1-F(\gamma))\right]}\right] \tag{24}$$

On the other hand, for large enough values of the interest rate, the solution is given by the system formed by equations

$$nk^h=\gamma\frac{1}{\nu}\left[r\left(\theta^c-1\right)+\left(\theta^c-1\right)\right] \tag{25}$$

and (24) and the equilibrium value of the fee at the cutoff point is zero, $\frac{q(\gamma)}{p}=0$.

We would like to stress two features of the solution. The first, is the comparison of the quadratic equation obtained now, (24), with the one of Section 3, obtained assuming market determined interest rates, equation (11), repeated here

$$\frac{n^2\phi}{(1-\phi n)}=\frac{(\theta^c-1)\Omega(\gamma)+r\left[\theta^c\Omega(\gamma)+\theta^d\left[\Omega(h(r)\gamma)-\Omega(\gamma)\right]+\theta^a\left[1-\Omega(h(r)\gamma)\right]\right]}{\left[1+k^d\left(F(h(r)\gamma)-F(\gamma)\right)+k^a\left(1-F(h(r)\gamma)\right)\right]}.$$

The right hand side of these equations summarizes, in each case, the interaction between the decisions regarding the cutoff point (γ), and the number of trips to the bank. In a standard Baumol-Tobin model, in which the cutoff is fixed, the only changes in the right hand side of this equation would come about changes in the interest rate. However, the inside money endogenous decision (γ), also changes the right hand side of that equation.

For instance, for the parametrization we adopted in Section 5, the range of variation of the right hand side of equation (11), goes from

$$\frac{r\theta^a}{1+k^a} = \frac{0.01}{1+0.049}r$$

when $\gamma \to 0, \delta \to 0$, (so to $\Omega(\gamma) = 0 = \Omega(h(r)\gamma)$, and $F(h(r)\gamma) = 0 = F(\gamma)$) to

$$(\theta^c - 1) + r\theta^c = 0.01 + 1.01r$$

when $\gamma \to \infty, \delta \to \infty$, (so to $\Omega(\gamma) = 1 = \Omega(h(r)\gamma)$, and $F(h(r)\gamma) = 1 = F(\gamma)$).

In this case, the difference between the two extremes is given by

$$0.01 + 1.01r - \frac{0.01}{1+0.049}r \simeq 0.01 + r$$

the size of the interest rate.

On the other hand, the range of variation in the right hand side of equation (24), is given by

$$\frac{r}{1+k} = \frac{r}{1+0.03}$$

when $\gamma \to 0$ (so to $\Omega(\gamma) = 0$, and $F(\gamma) = 0$) to

$$\frac{r\theta^c + (\theta^c - 1)}{1} = 1.01r + 0.01$$

when $\gamma \to \infty$,(so to $\Omega(\gamma) = 1$ and $F(\gamma) = 1$). Thus, in this case, the difference between the two extremes is given by

$$1.01r + 0.01 - \frac{r}{1+0.03} \simeq 0.01r + 0.01$$

or $1/100$ of the interest rate.

What these numbers suggest is that the feedback between the portfolio decision and the number of trips to the banks will be quantitatively very small during the regulation Q period, compared with the ones of the market determined interest rates period. As we explained in Section 3, it was this feedback the reduced the interest rate elasticity when interest rates are freely determined. As the quantitative impact of this feedback is reduced, we expect the interest elasticity of real money balances to be higher during the regulation Q period.[21]

The second feature, is related to the other equation that determines the solution, given by (23) for very low values of the interest rate, and by (25) for interest rates higher than the threshold given by (22). In the first case, the equation is the same as in Section 3, so the solution is similar to that one in Section 3, except for the discussion above, since the quadratic equation is much flatter than in Section 3.

However, once the interest rate reaches the threshold, the straight line is now given by (25). This equation is much less sensitive to increases in r, since it is multiplied by $(\theta^c - 1)$ instead of $\left(\theta^c - \theta^d\right)$ which is way smaller.[22] Thus, this equation is much less sensitive to changes in the interest rate. Therefore, increases in r beyond the threshold (22), increase n, as before, but now they increase γ. The general solution is depicted in Figure 6.1.[23]

Initially, when the interest rate goes from low to medium, the system behaves similar to the one in Section 3, the number of trips to the Bank goes up (real balances go down) and the ratio of cash over money goes down. The system is not the same as before, since the quadratic equation is different from the one in Section 3 - the concave curve is much flatter than before. As discussed above, this system delivers a larger effect on real money balances following an increase in the nominal interest

[21] Note that goes in the right direction, relative to the data showed in Figure 5.8.

[22] For our calibration, $\theta^c - \theta^d = 0.91$, while $\theta^c - 1 = 0.01$.

[23] We used the parameter values of the calibrationbelow, and interst rates of 1%, 3% and 5%.

rate.

Once the interest rate reaches the threshold, further increases of the interest rate from medium to high imply an increase of n, but larger than before, since the now the threshold γ increases. This also pushes the interest elasticity of real money balances up during regulation Q - though the quantitative impact is very small, since the concave curve is very flat. In addition, note from Figure 6.1, that the relationship between r and γ is U-shaped. As it turns out, the U-shape pattern is also exhibited in the data between 1935 and 1981 - the period when regulation Q had full applicability.[24]

6.1 Special Cases

We now briefly consider two special cases; namely, the case in which all fixed cots are born by banks and the case in which there is no cost of using cash, so $\theta^c = 1$.

When $k^h \to 0$, then threshold is satisfied for all finite r, so that in equilibrium, $q(\gamma) > 0$ always. The reason is that, absent any fixed cost by households, there would be no reason to hold cash if the fee were zero. Thus, competition across banks will push the threshold γ to a low enough value such that the fee at the margin will always be positive, no matter how large r is.

On the other hand, when $\theta^c \to 1$, the threshold is never satisfied, and equation (25) implies that $\gamma \to \infty$. In this case, as there is no advantage of holding deposits, only cash is used in equilibrium.

[24] The U-shape pattern in the data, depicted in Figure 6.3, is likely affected by potentailly lasting effects of the the bank panics of the 30's on one hand and the war in the other. Thus, we do not want to push this too far. However, the model does imply that for higher interst rates, there is an increasing relationship between r and γ, as it is observed in the data from 1950 to 1980.

6.2 Results

We have already calibrated all parameters of the model. However, to evaluate its performance during the regulation Q period, we need to assign values to the additional parameter, $k^h = k - k^b$. In other words, we need to split the fixed costs of making transactions (0.03 in our calibration) between banks and households, a distinction that is immaterial with market determined interest rates.

We let k^h be such that the U-shape implied by γ in the theory matches the minimum in the U-shape curve of the data. The results for real money balances are depicted in Figure 6.2. As argued before, the curve during the regulation Q period is steeper than the curve when deposits can pay interest rates. However, the model implies an elasticity that is higher than the one suggested by the data: our theory underpredicts money balances during the Regulation Q period for interest rates that are higher than 1%. As before, this curve is very robust to variations in the values used for calibration.

Finally, in Figure 6.3 we plot the theoretical values for the cash to money ratio for the regulation Q period, together with the data. As it can be seen, the model strongly over-predicts the level of the ratio. This feature, as it was the case in Section 3, is non-robust to changes in the values used for the calibration. For instance, changing the parameter θ^c, or k^d substantially change the level of the cash to money ratio, while leaving essentially unchanged the slope of real money balances.

As before, the behavior of each of the components is very sensitive to the banking technology in another dimension: small changes in the calibration target imply big changes in Figure 6.3. For instance, assuming that the GDP loss in trips to the bank by households (ϕn) is 1.5% of output, instead of 1%, improves substantially the cash to money ratio picture. On the other hand, the behavior of the real money balances is barely affected. We do not pursue this idea further. The reason why we chose to calibrate the parameters so that total resources spent on fixed cost of using deposits is 1% of GDP and total transaction costs of going to the bank is 1%, is not because we do have estimates of those numbers - we do not. It is because, with enough effort in empirical analysis, one could obtain those aggregate estimates. And those are the right aggregate numbers to discipline the values for k^d, k^a and ϕ.

In summary, our model of banking competition during regulation Q goes in the right direction qualitatively: it implies that the interest rate elasticity should be higher under this regulation. However, the overpredicts the increase in the interest rate elasticity during the period. In addition, it misses the level of cash to deposits ratio, though it can explain the increase in the relative use of cash during the increase in interest rates that occurred from 1950 to 1980. As before, the implications regarding the level of the ratio of cash to deposits is very sensitive to small changes in parameters we cannot precisely calibrate. On the other hand the behavior of the aggregate is very robust to those changes.

The data for M1 over GDP during the regulation Q period lies between the two theoretical models considered, one with market determined interest rates and the other with an effective cap on interest rates paid by banks at zero. The evidence on M1 seems to suggest that, somehow, banks could partially circumvent the ceiling on interest rates and transfer to depositors a fraction of the return they obtained with their deposits.

7. CONCLUSIONS

We view the new aggregate NewM1 as a step toward doing for M1 what Motley (1988) and Poole (1991) did for M2 with their new aggregate MZM. Both are attempts to define redefine monetary aggregates by the *functions* they have in the payments system rather than by the institutions whose liabilities they are. Roughly speaking, the deposit component of NewM1 includes and is limited to deposits you can write checks on, on paper or electronically, a measure of the same thing that M1 measured in the past. The new aggregate does about as well on low and medium frequencies over the period 1915-2013 as M1 did from 1915 to 1990, as about as poorly on high frequencies.

Our application of the Freeman/Kydland model of the components of these broader aggregates, reported in Sections 3 and 6, is frankly exploratory. The whole question of the substitutability among various means of payment is captured by the distribution of transaction sizes. We have just begun to think through and test the many possibilities.

Finally, since we are writing in 2012, we should emphasize that the analysis is all based on theoretical steady states. We are trying to get the quantity theory of money back to where it seemed to be in 1980, but after all the older theories were as silent on financial crises as is this one is.

List of References

Alvarez, F. and F. Lippi. "Financial innovation and the transactions demand for cash", Econometrica, 2009

Baumol, William J. "The Transactions Demand for Cash: An Inventory Theoretic Approach." *Quarterly Journal of Economics*, 66 (1952): 545-556.

Freeman, Scott, and Finn E. Kydland. "Monetary Aggregates and Output." *American Economic Review*, 90 (2000): 1125-1135.

Ireland, Peter N. "On the Welfare Cost of Inflation and the Recent Behavior of Money Demand." *American Economic Review*, 99 (2009): 1040-1052.

Lucas, Robert E., Jr. "Inflation and Welfare," *Econometrica,* 68 (2000): 247-274.

Meltzer, Allan H. "The Demand for Money: The Evidence from Time Series," Journal of Political Economy, 71 (1963): 219-246.

Motley, Brian. "Should M2 be Redefined?". Federal Reserve Bank of San Francisco *Economic Review*, (1988), 33-51.

Philippon, Thomas. "The Evolution of the U.S. Financial Industry from 1860 through 2007." NYU Working Paper (2008).

Poole, William. Statement Before the Subcommittee on Domestic Monetary Policy of the Committee on Banking, Finance and Urban Affairs. U.S. House of Representatives, (1991).

Prescott, Edward C. "A Multiple Means of Payment Model," in William A. Barnett and Kenneth N. Singleton, eds., *New Approaches to Monetary Economics,* Cambridge University Press (1987): 42-45.

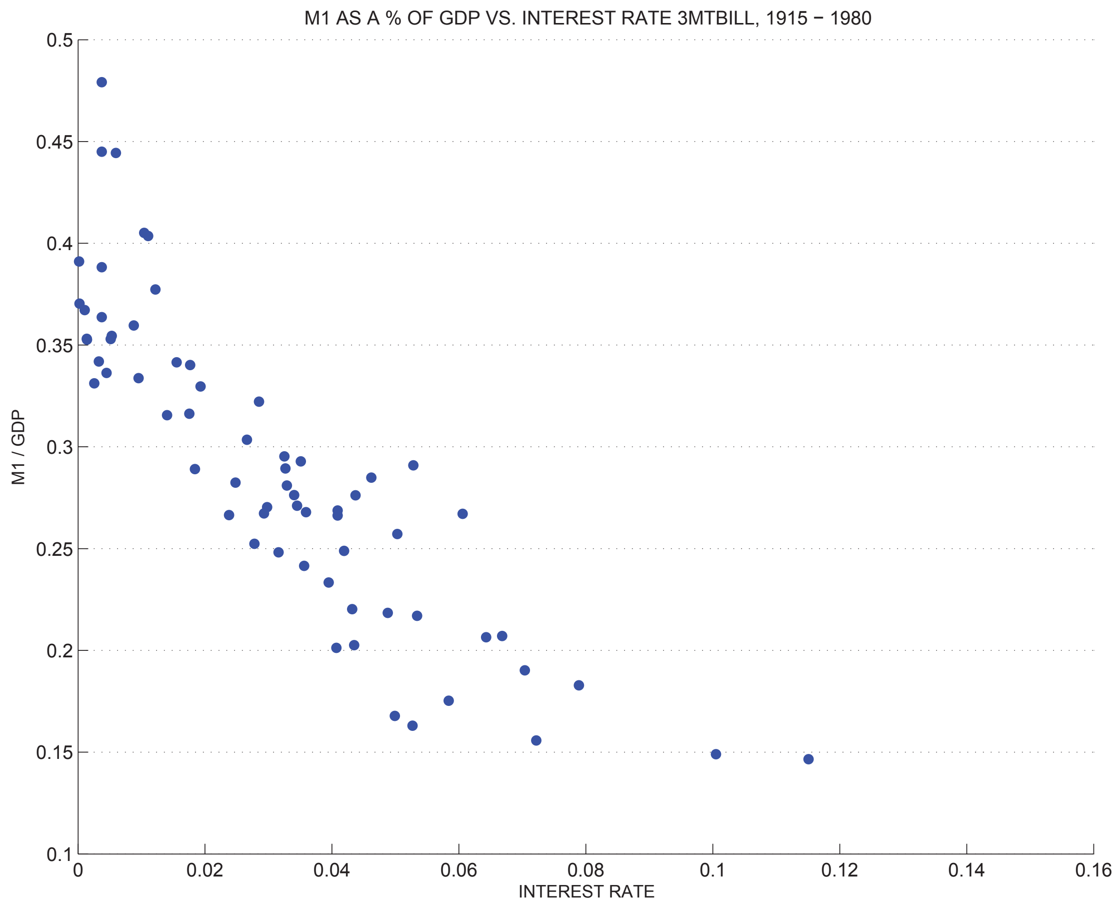
Reynard, Samuel. "Financial Participation and the Apparent Instability of Money Demand." *Journal of Monetary Economics*, 54 (2004): 1297-1317.

Teles, P. and R. Zhou, "A Stable Money Demand: Looking for the Right Monetary Aggregate." Federal Reserve Bank of Chicago *Economic Perspectives,* (2005): 50-63.

Tobin, J. "The Interest Elasticity of Transactions Demand for Cash." *Review of*

Economics and Statistics, 38 (1956): 241-247.

Figure 2.1



M1 AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 1980
M1 / GDP
INTEREST RATE
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.2

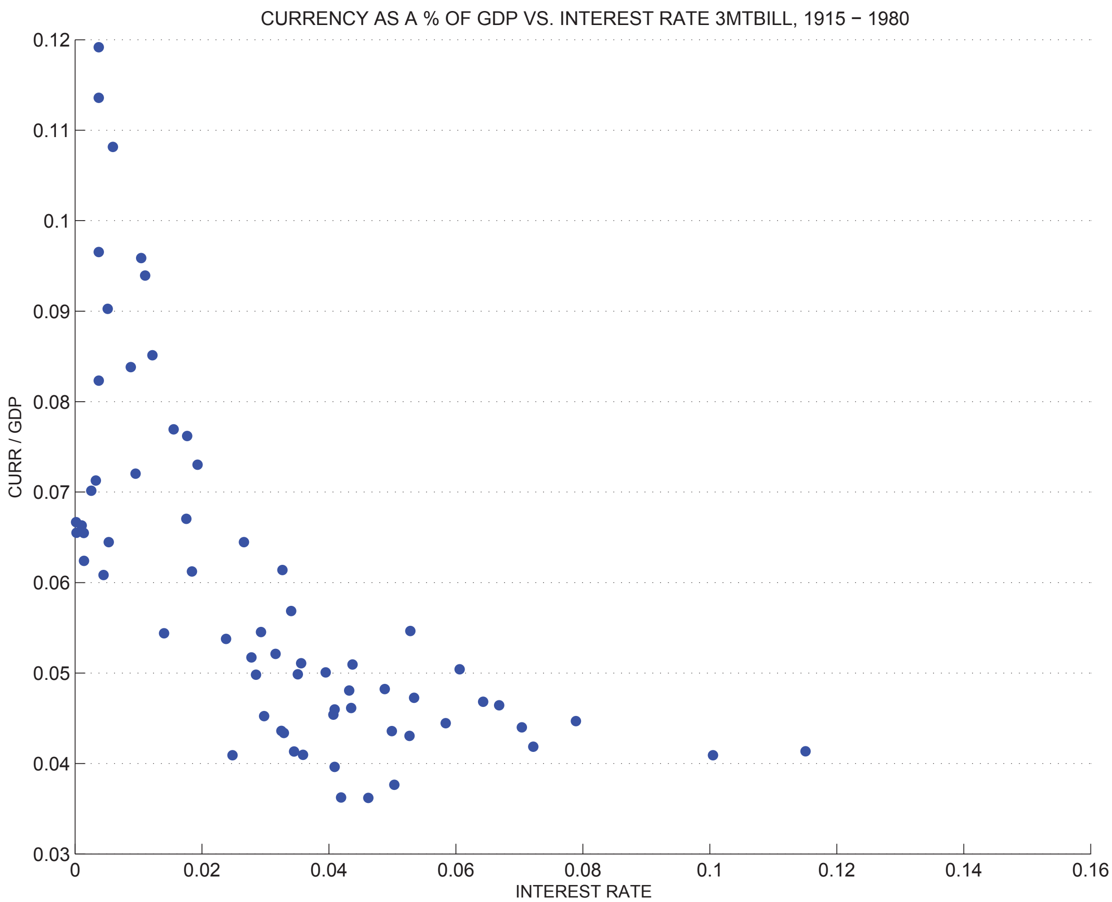

CURRENCY AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 1980
CURR / GDP
INTEREST RATE
0.12
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.3

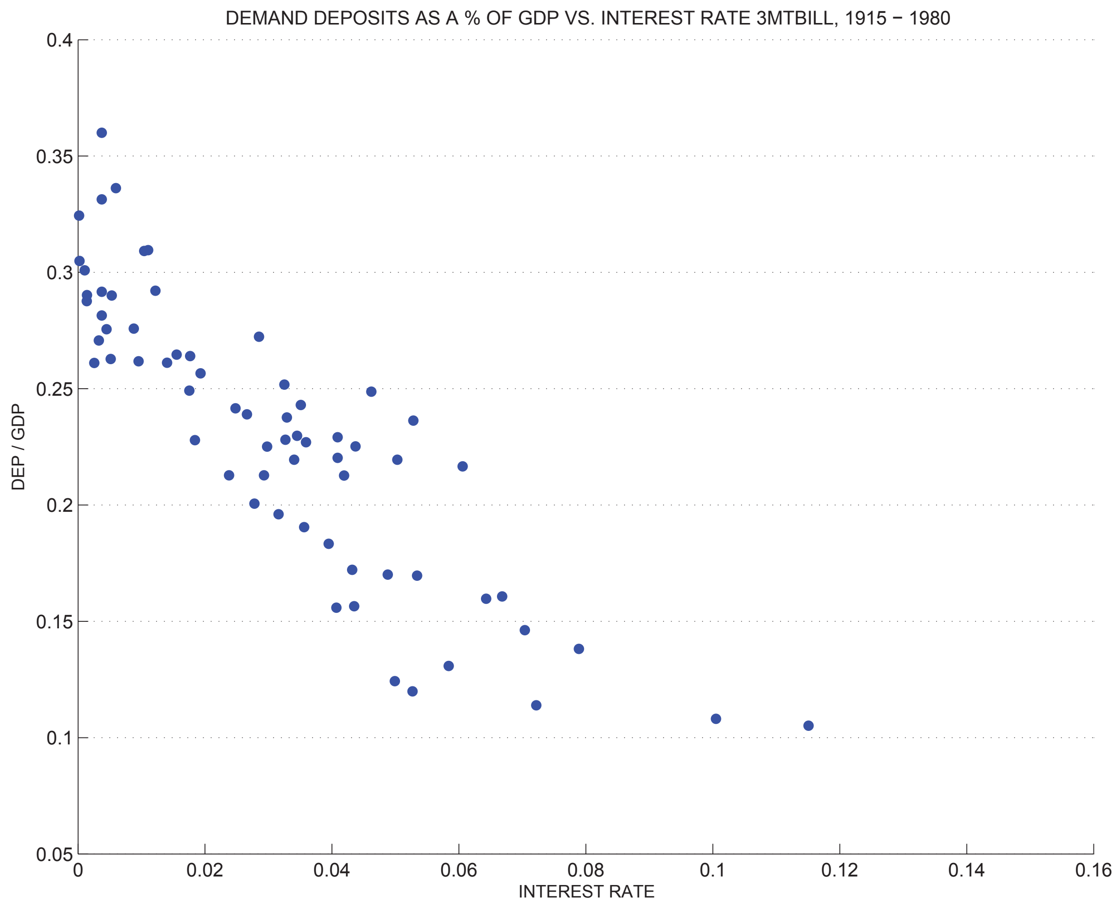
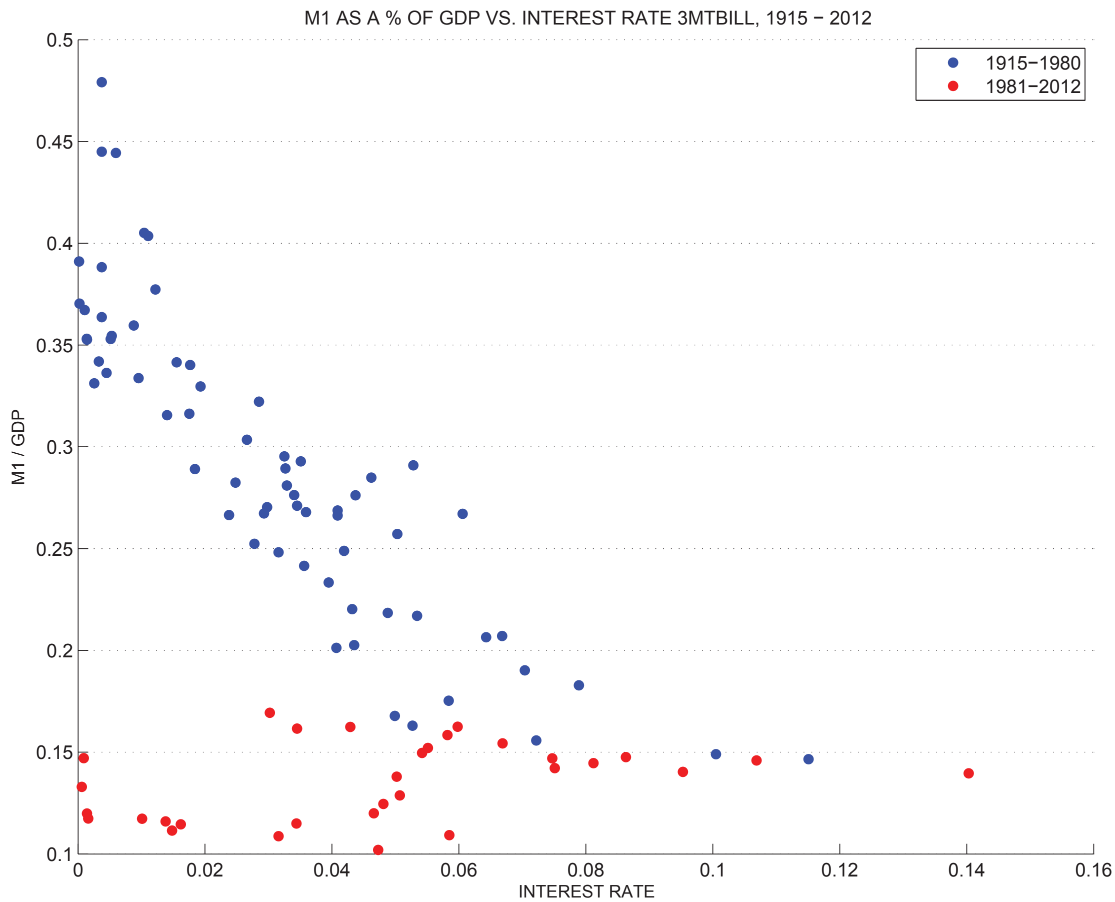

DEMAND DEPOSITS AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 1980
DEP / GDP
INTEREST RATE
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.4



M1 AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 2012
1915−1980
1981−2012
M1 / GDP
INTEREST RATE
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.5

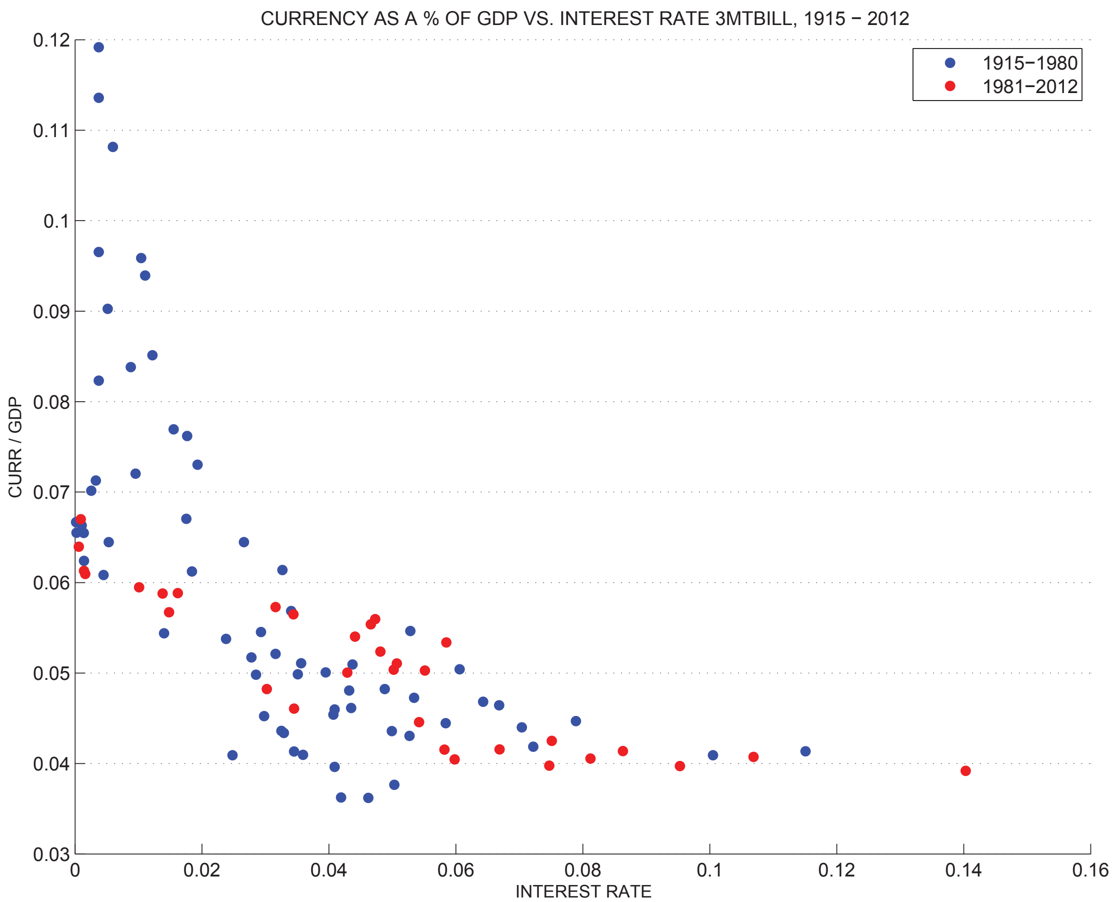

CURRENCY AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 2012
1915−1980
1981−2012
CURR / GDP
INTEREST RATE
0.12
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.6

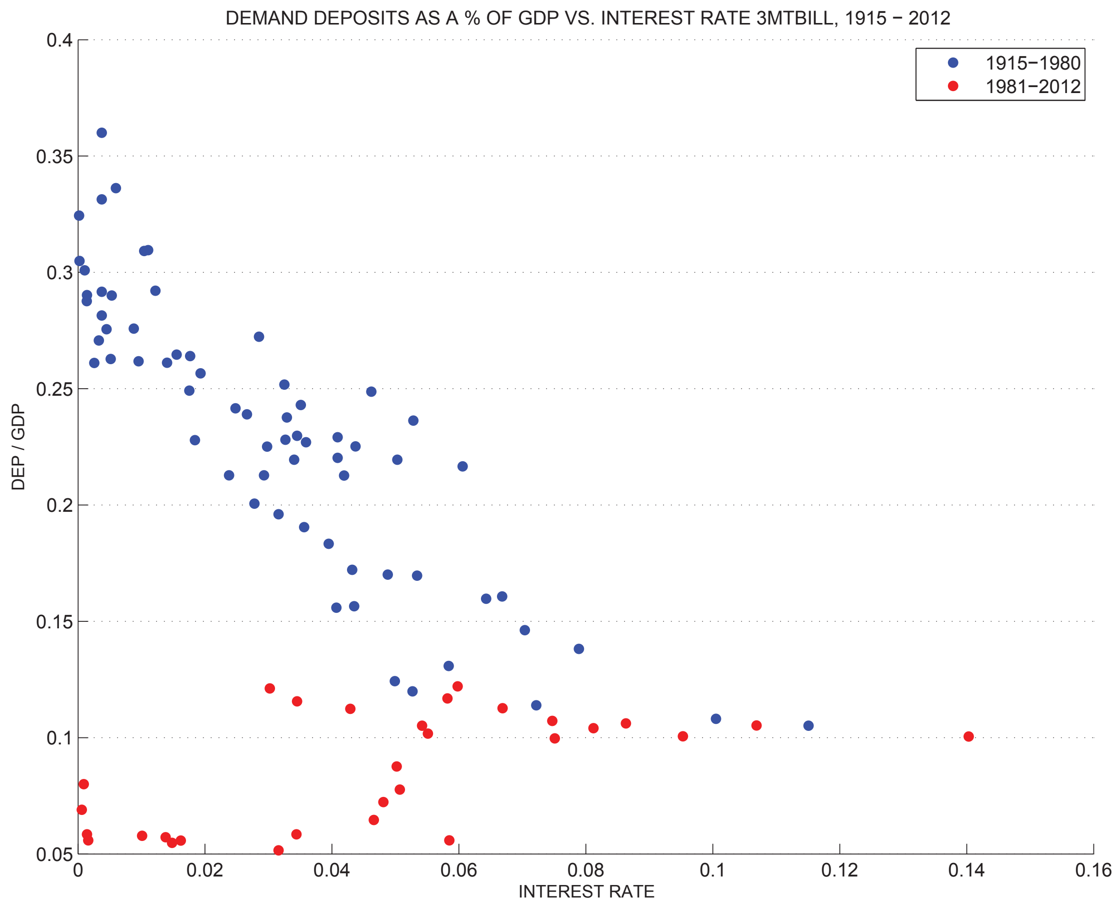

DEMAND DEPOSITS AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1915 − 2012
1915−1980
1981−2012
DEP / GDP
INTEREST RATE
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 2.7

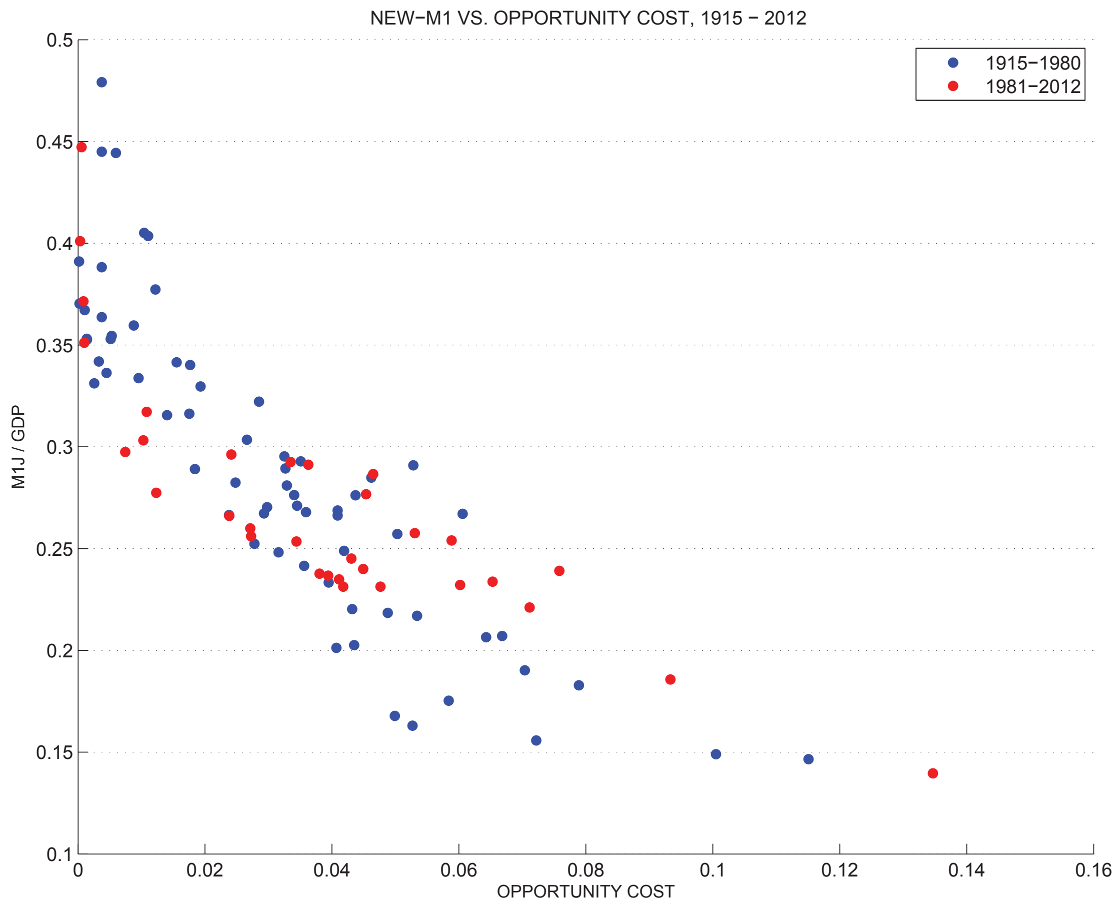

NEW−M1 VS. OPPORTUNITY COST, 1915 − 2012
1915−1980
1981−2012
M1J / GDP
OPPORTUNITY COST
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16

Figure 3.1

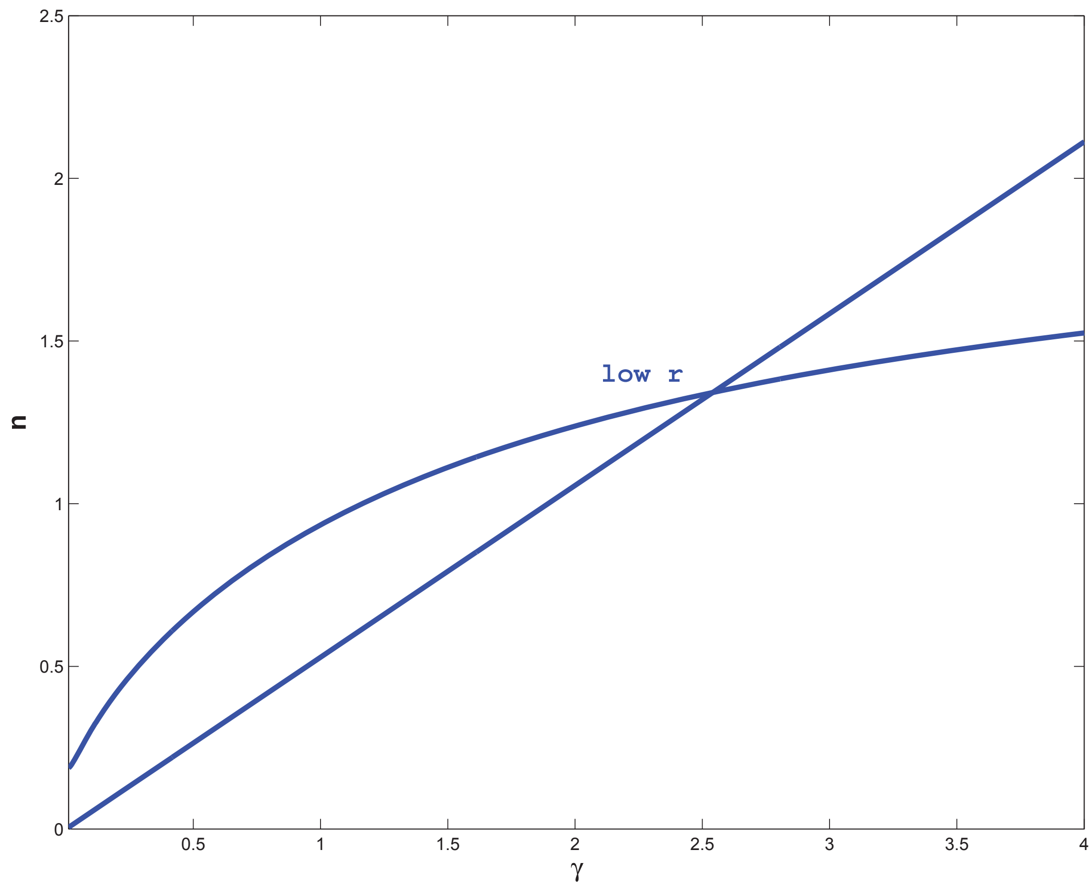


Figure 3.2

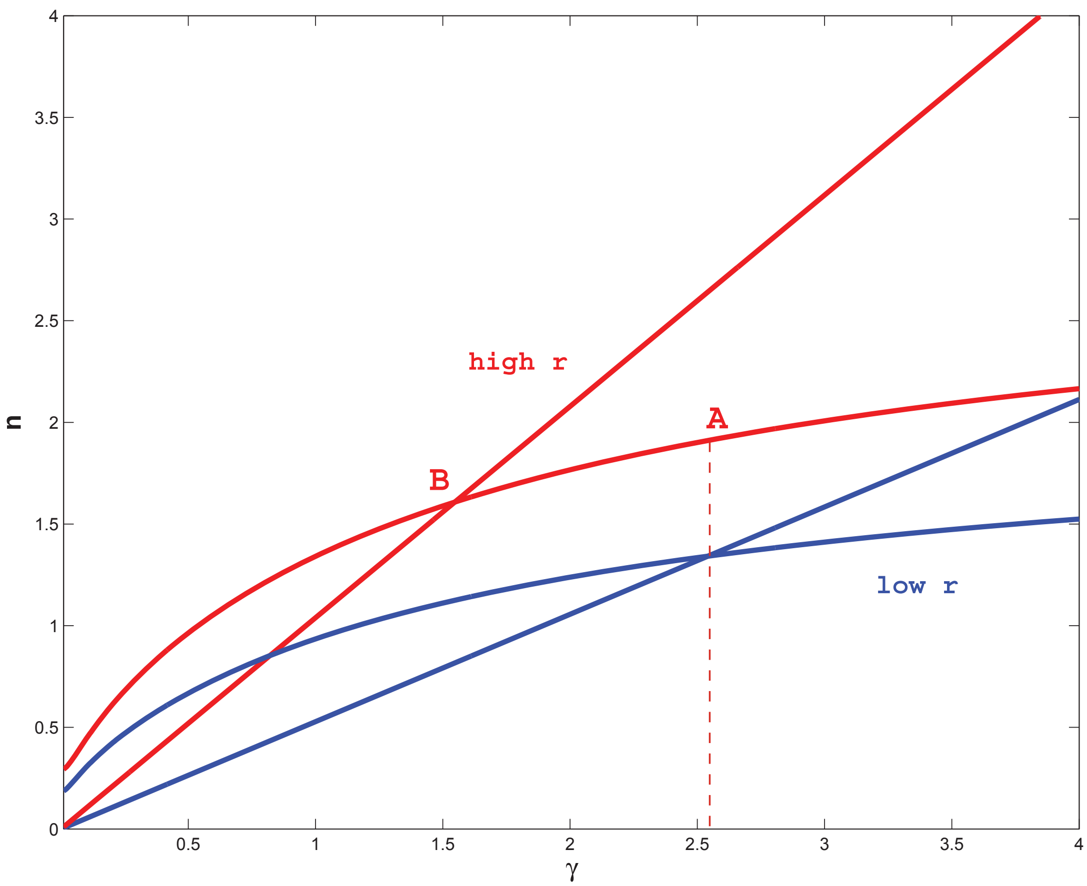
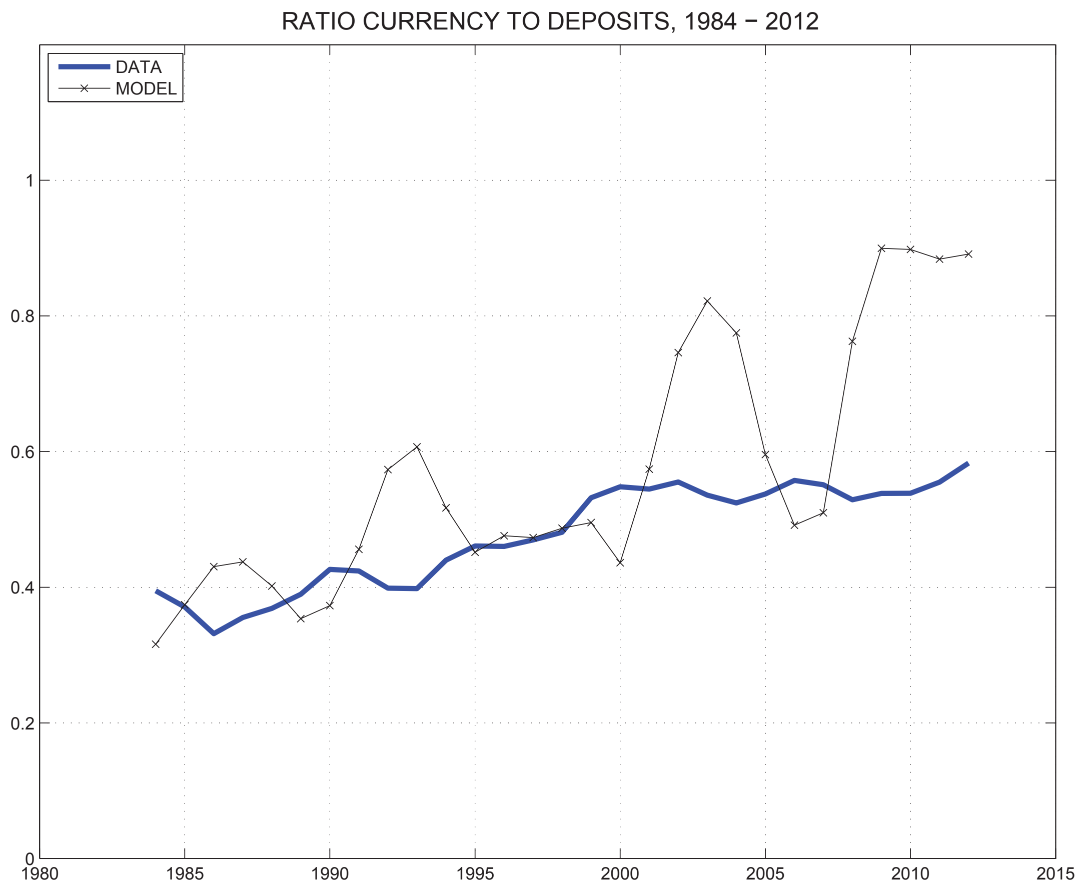


Figure 5.1



RATIO CURRENCY TO DEPOSITS, 1984 − 2012
DATA
MODEL
1
0.8
0.6
0.4
0.2
0
1980
1985
1990
1995
2000
2005
2010
2015

Figure 5.2

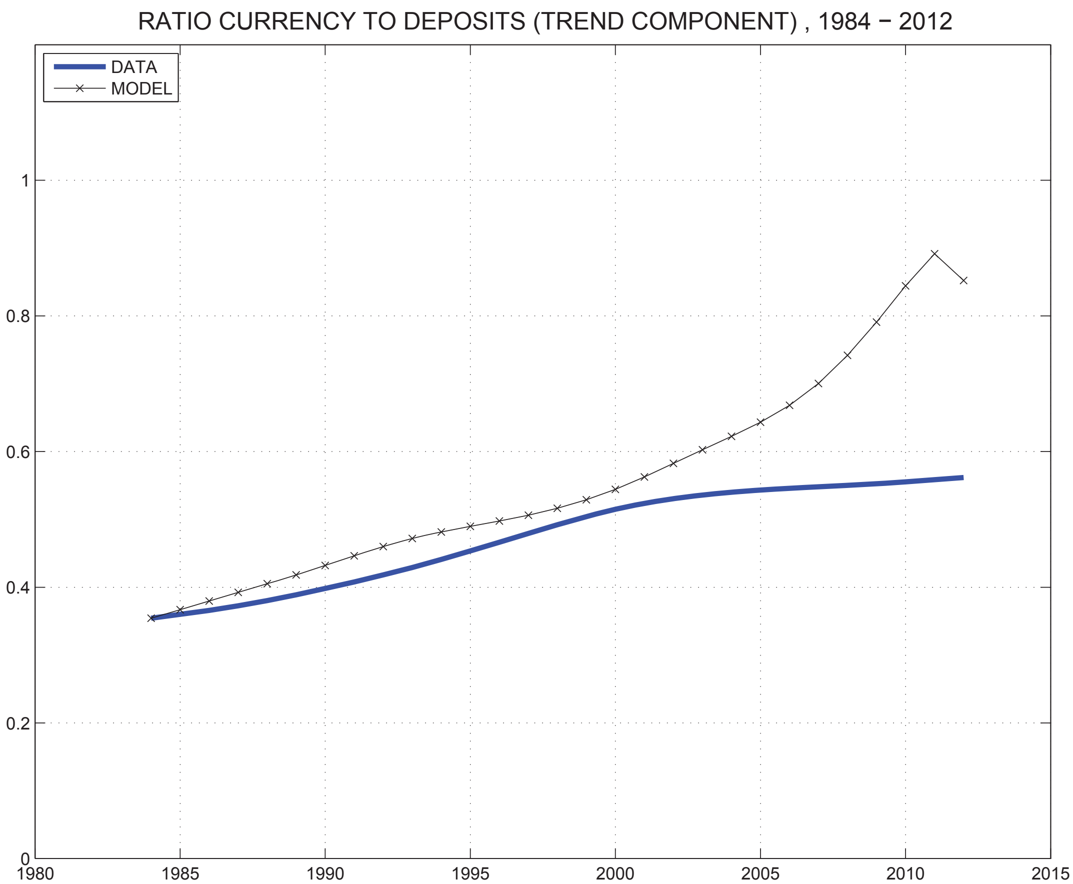

RATIO CURRENCY TO DEPOSITS (TREND COMPONENT) , 1984 − 2012
DATA
MODEL
0
0.2
0.4
0.6
0.8
1
1980
1985
1990
1995
2000
2005
2010
2015

Figure 5.3

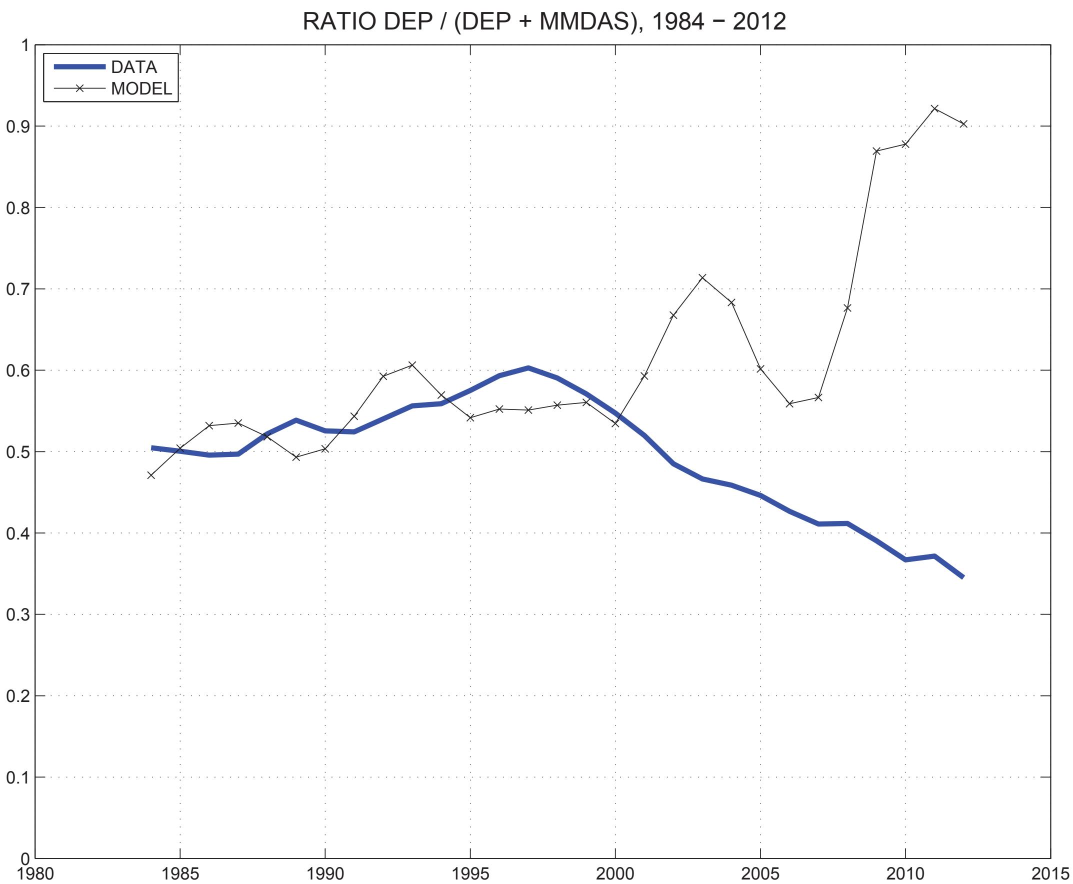
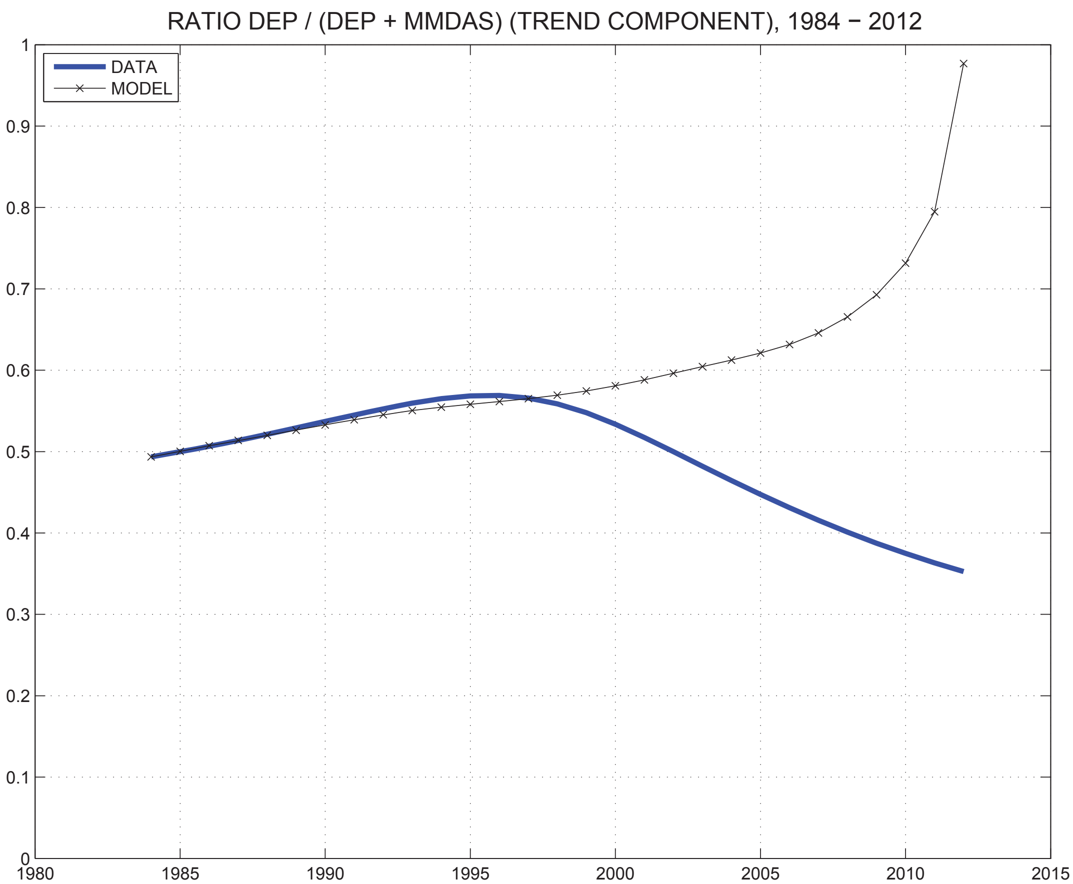

RATIO DEP / (DEP + MMDAS), 1984 − 2012
DATA
MODEL
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1980
1985
1990
1995
2000
2005
2010
2015

Figure 5.4



RATIO DEP / (DEP + MMDAS) (TREND COMPONENT), 1984 − 2012
DATA
MODEL
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1980
1985
1990
1995
2000
2005
2010
2015

Figure 5.5

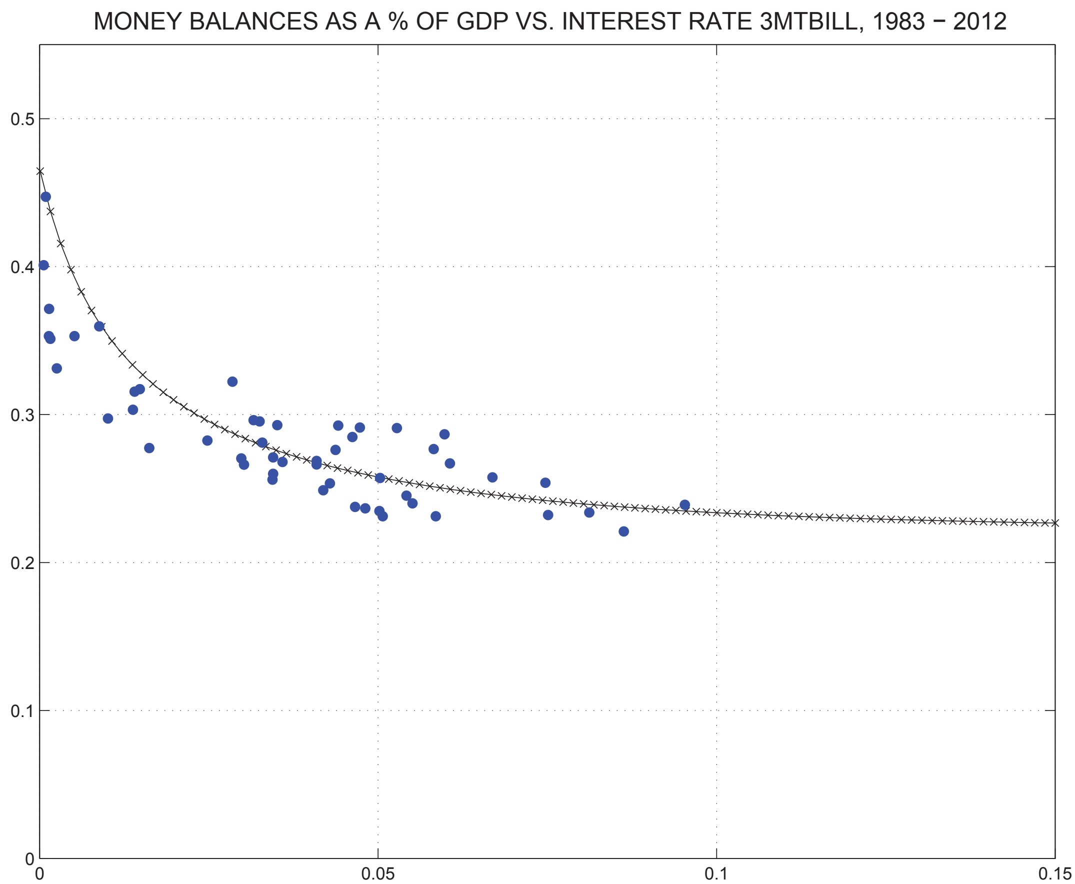
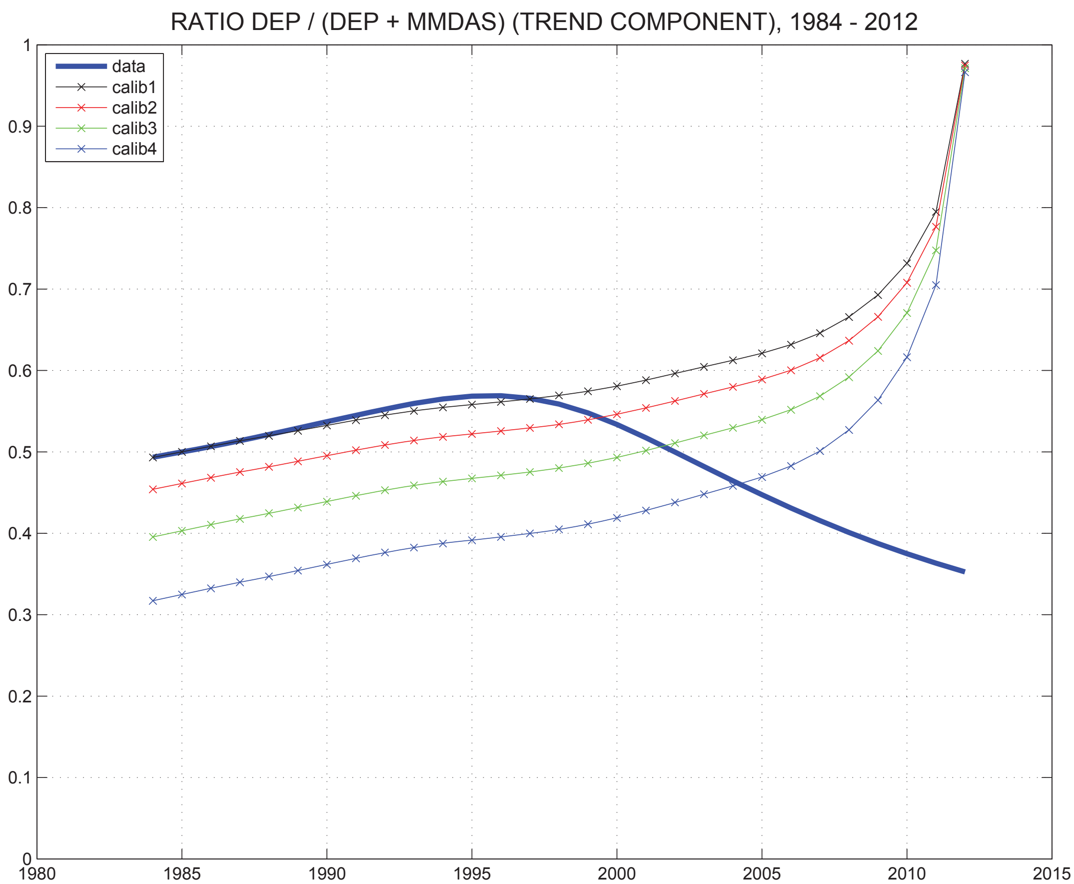

MONEY BALANCES AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1983 − 2012
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15

Figure 5.6



RATIO DEP / (DEP + MMDAS) (TREND COMPONENT), 1984 - 2012
data
calib1
calib2
calib3
calib4
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1980
1985
1990
1995
2000
2005
2010
2015

Figure 5.7

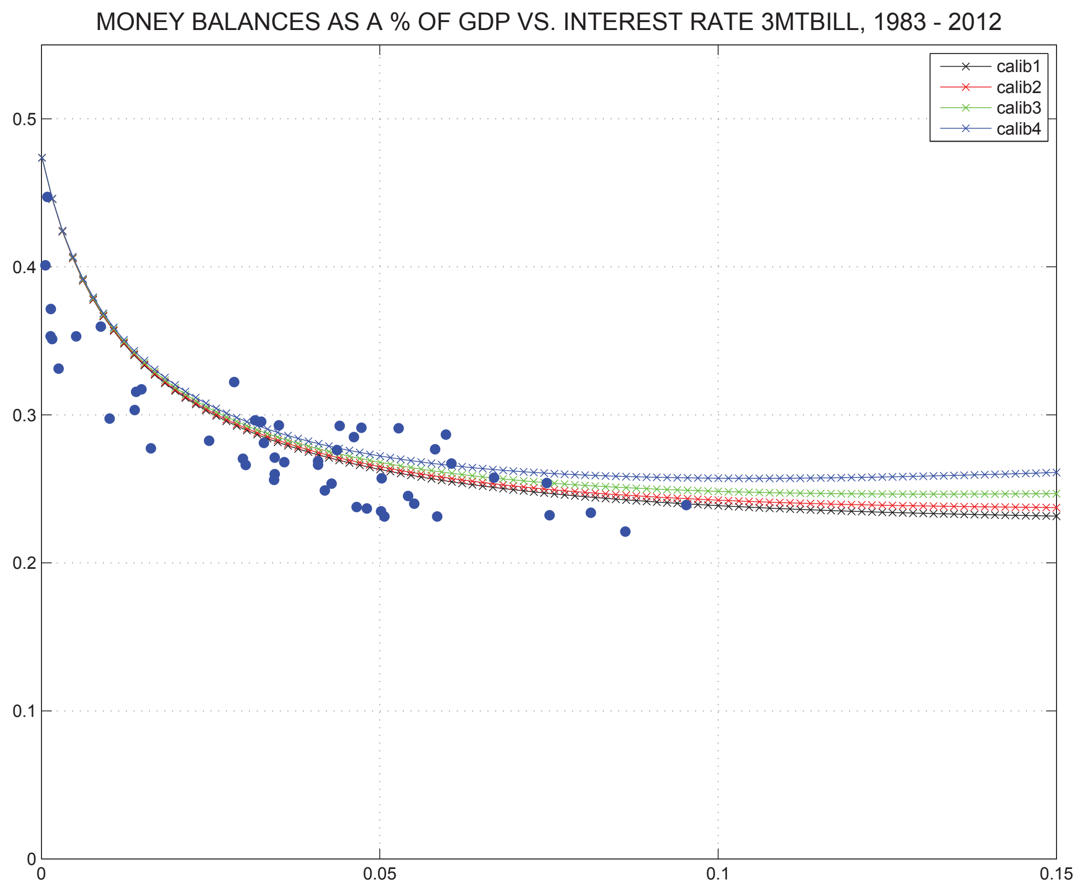
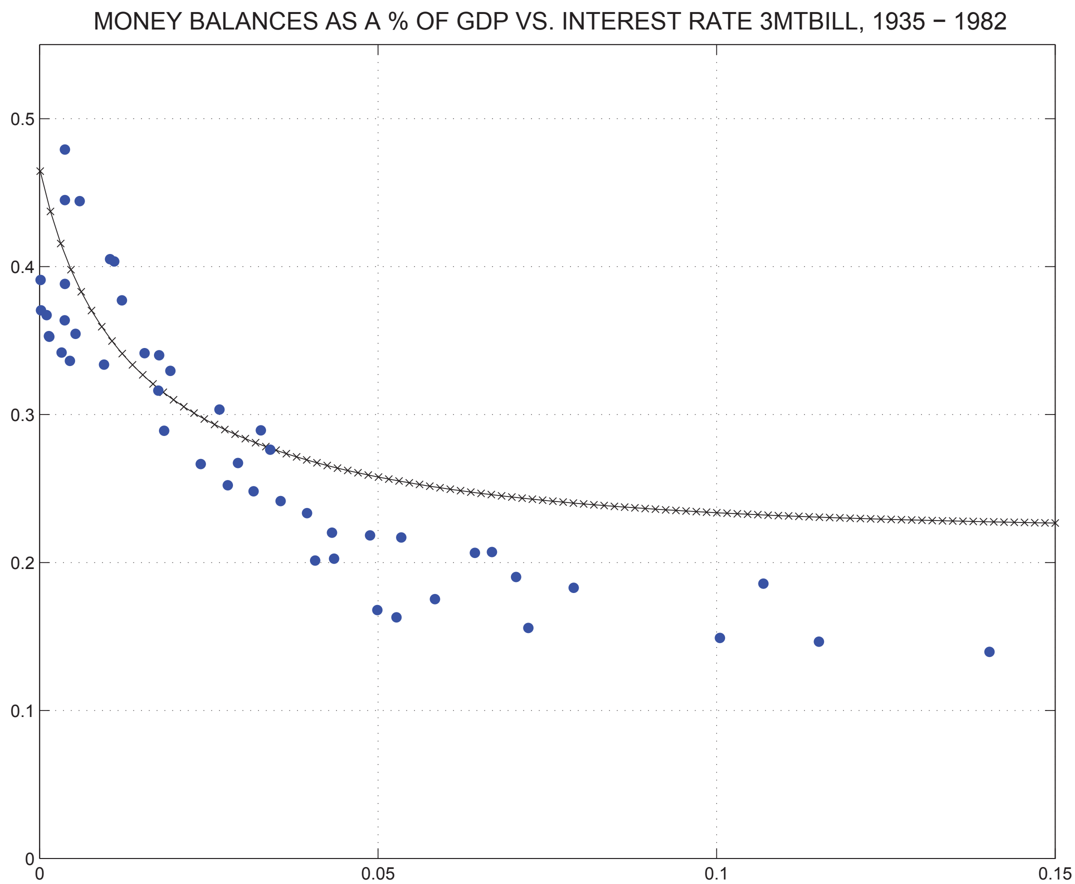

MONEY BALANCES AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1983 - 2012
calib1
calib2
calib3
calib4
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15

Figure 5.8



MONEY BALANCES AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1935 − 1982
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15

Figure 6.1

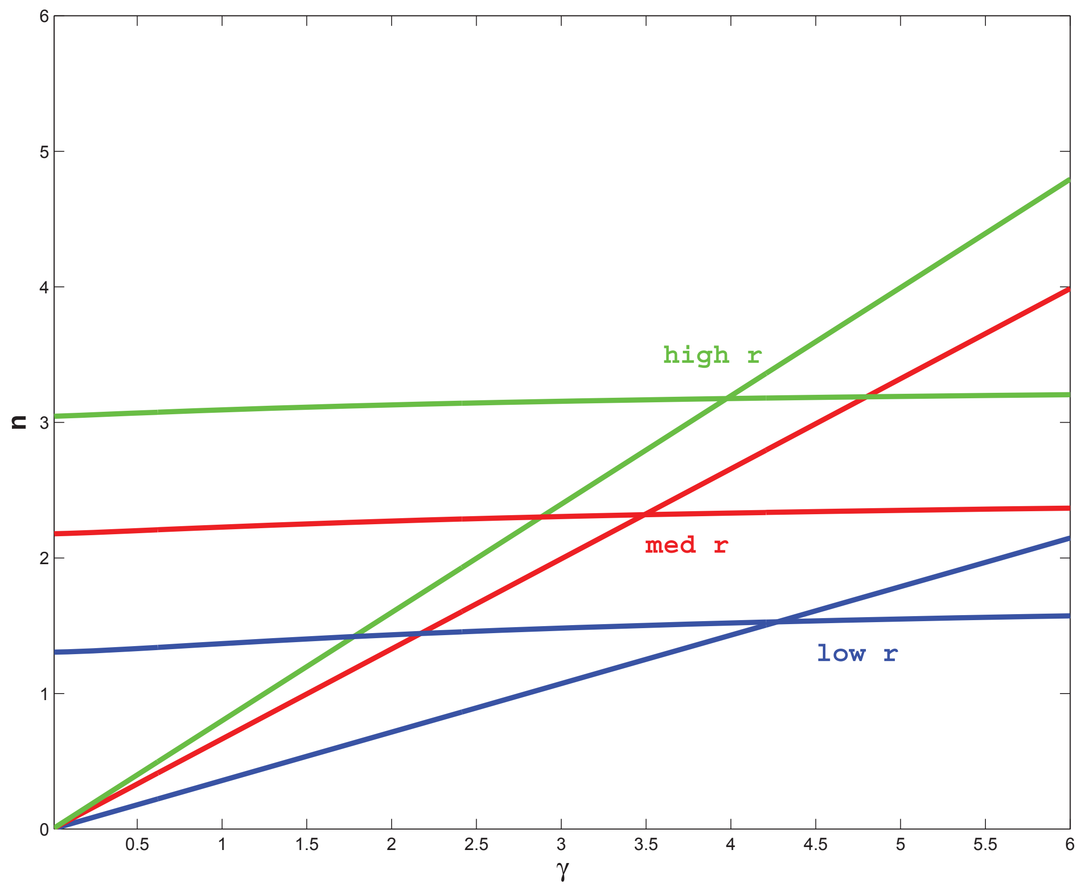


Figure 6.2

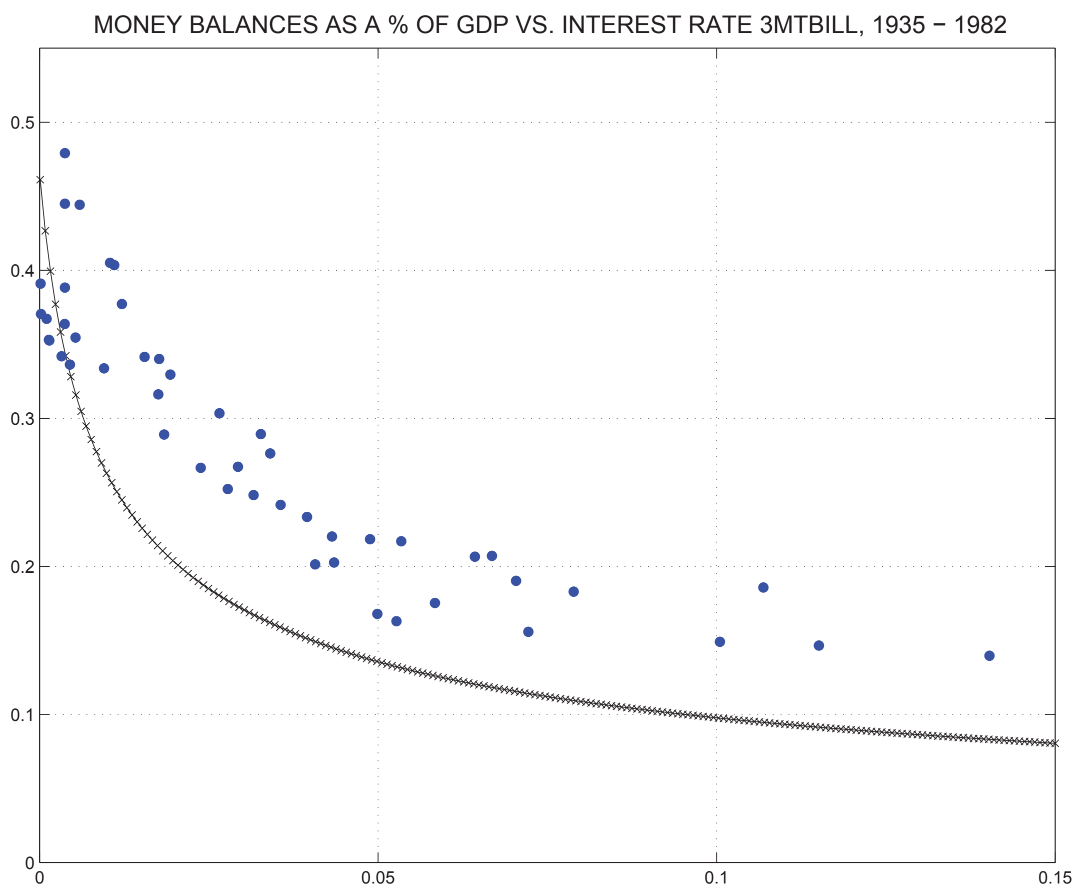
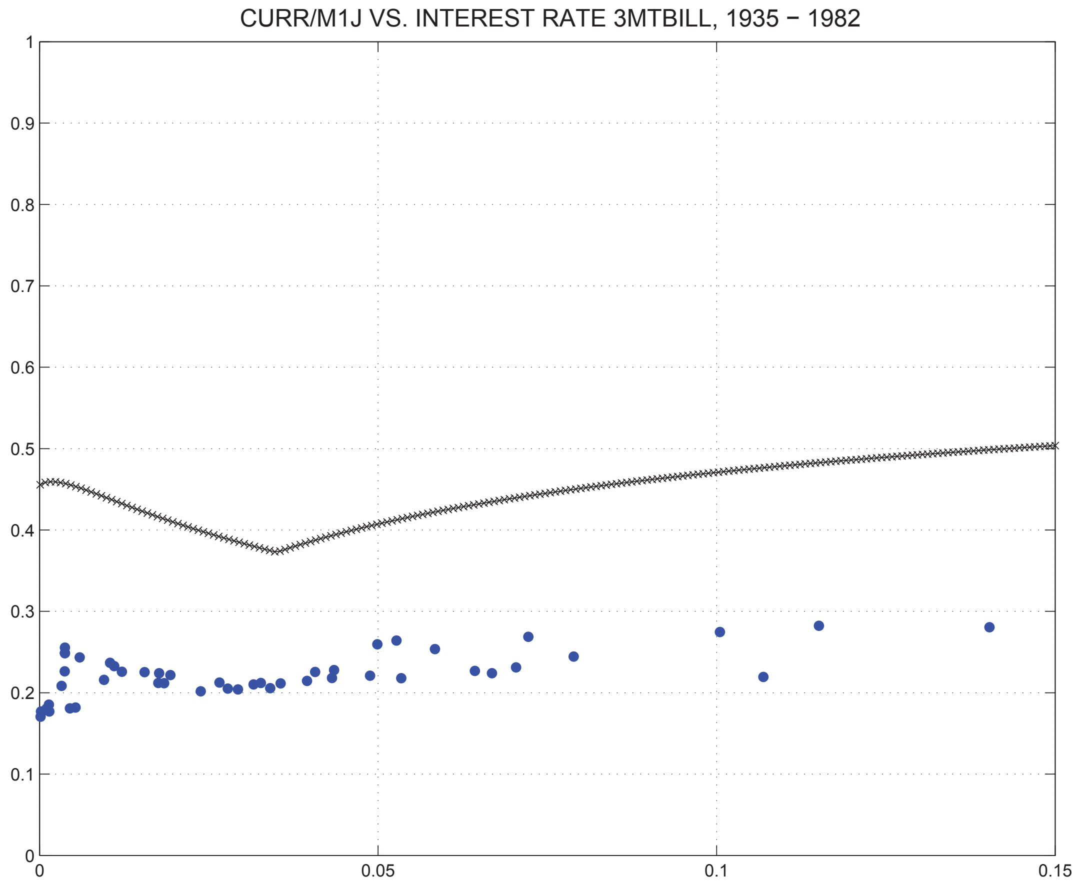

MONEY BALANCES AS A % OF GDP VS. INTEREST RATE 3MTBILL, 1935 − 1982
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15

Figure 6.3


CURR/M1J VS. INTEREST RATE 3MTBILL, 1935 − 1982
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15