MACROFOUNDATIONS FOR A (NEAR) 2% INFLATION TARGET

Christophe Faugère

June 2010

*I thank Richard Anderson and Peter Ireland for their valuable comments, Don Dutkowsky for graciously answering my inquiries regarding his (and his co-authors’) database, and finally the participants at the fall 2009 UAlbany Economics seminar for their suggestions. Any remaining errors are of course mine.

*Department of Finance and Center for Institutional Investment Management, University at Albany, Albany, NY 12222. E-mail: cfaugere@uamail.albany.edu.
MACROFOUNDATIONS FOR A (NEAR) 2% INFLATION TARGET

Abstract. Economists have argued that a long-term inflation target near 2% is optimal (Summers, 1991; Fischer, 1996; Goodfriend, 2002; Coenen et al., 2003; Bernanke, 2003). However, these arguments are really about why a low positive inflation rate is ideal to avoid a deflationary trap, not explaining why the specific value of 2% (or a value near it) happens to be the optimal long-run inflation rate. In line with the transaction motive literature (Baumol, 1952 and Tobin, 1956), I postulate that financial innovations generate transactional cost savings by comparison to barter. The overall trend in productivity growth affecting financial transactions and the advent of credit cards in the world economy are crucial factors. I derive the optimal velocity of money, which depends on real GDP/capita and the net return on depository institutions’ assets. As long as progress is on average biased towards new forms of money, the velocity of money will grow at a pace slower than long-term real GDP/capita growth; i.e. less than 2%. Along with a parameter representing the type of bias in the technical progress affecting transactions, the depository institutions’ overall mean leverage ratio also appears as a key parameter in the long-run equilibrium equation describing the behavior of the velocity of narrow money (M1, M1RS and M1S). I use Johansen’s (1988, 1991 and 1995) VECM approach for the U.S. from 1959-2007 and find good support for the model. I show that a ‘naïve’ Friedman k-percent monetary rule that aims at growing the money supply at the same rate as real GDP naturally leads to a rate of inflation equal to the rate of velocity growth. Hence, setting an inflation target near but below 2% makes economic sense. In spite of previously held beliefs, a money growth objective is compatible with an interest-targeting objective; i.e. a derived Taylor (1993) type rule. A Taylor rule that embeds the optimal inflation target defined here is more flexible to account for possible changes in velocity vs. a pure money growth rule.

Keywords: Inflation target, velocity of narrow money, M1, M1RS, M1S, real GDP per capita growth, barter, financial leverage.

JEL: E40
1. Introduction

Although current Fed chairman Bernanke has long been a fervent advocate of inflation targeting (Bernanke et al., 1999), and although, the U.S. Federal Reserve has been setting implicit inflation targets off and on over the past two decades, interest rate targeting has been and still is the primary tool of U.S. monetary policy. Economists are well aware of the limitations associated with both inflation and interest rate targeting (Kozicki, 1999; Benhabib et al., 2001; Carlstrom and Fuerst, 2003; Sims, 2004; McCallum, 2006). Yet, there is clearly a scope for both methods to serve as complementary monetary policy tools. An instance of this conjunction is the U.S. monetary policy of the late 1980s, which is well characterized by John Taylor’s interest rate rule (1993). As is sometimes overlooked, Taylor’s (1993) rule contains both a 2% interest rate target and a 2% inflation target in its formulation.

Inflation targeting was first adopted as a primary policy instrument by the Reserve Bank of New Zealand in 1988. Other G-7 central banks followed suite in the early 1990s. Since then a consensus has emerged among central bankers and economists that a narrow inflation range around 2% is optimal (see Table 1). The extant literature essentially demonstrates that a low positive inflation rate is ideal to avoid a deflationary trap. It stops short of providing any micro- or macro- foundation for why 2% happens to be the “magic” number. For example, in a speech at the St. Louis Fed 28th Annual Policy

---

1 The U.S. Fed has targeted short-term interest rates intermittently since the 1920s under the Riefler-Burgess doctrine. Prolonged periods of fed fund targeting occurred in the 1950s and 1960s (Meltzer, 2003 and 2009). The 1980’s is often designated as the beginning of a more systematic interest rate targeting policy in the U.S. (Bernanke and Mishkin, 1992; Thornton, 2005).

2 In a classic article, Taylor (1993) shows that the behavior of short-term interest rates from 1987 to 1992 follows a simple feedback rule based on achieving a 2% real interest rate target, correcting for deviations of inflation from a 2% target, as well as percentage deviations of real GDP from trend. Goodfriend (2005) argues that the Fed effectively practiced inflation targeting during the Greenspan years, but that interpretation has not been endorsed by Fed officials (Kohn, 2005). The presence of a 2% inflation target in Taylor’s (1993) rule is indicative that Goodfriend’s view may have some validity.

3 It is important to point out that 2% is a focal point in the literature. Many articles either cite 2% explicitly or give a desired narrow range around 2%. Summers (1991) asserts that the optimal long-run inflation rate is between 2 and 3%. However, his argument is brief and sketchy and does not justify why this specific range is best. Fischer (1996) lists a series of informal arguments centered on the Phillips curve and the difficulties of dealing with a zero inflation rate for stimulating the economy during slowdown periods. He states that “These arguments point to a target inflation rate in the 1 to 3 percent range; more specifically, they suggest that inflation should be targeted at about 2 percent, to stay within a range of 1 to 3 percent per year. This is in practice what most central banks mean by price stability; it is also a target that most G-7 central banks have already attained.” Goodfriend (2002) states that “Strictly targeting inflation between 1 and 2 percent could firmly anchor expected inflation and still give a central bank leeway to push the real
Conference, Bernanke (2003) introduces the concept of optimal long-run inflation rate. The speech’s main reference is an article by Coenen et al. (2003) who support a 2% target in order to keep deflation at bay. Coenen et al. simulate the impact of various inflation targets on the means and variances of inflation and output. They find that lowering the target inflation rate down to 2% has little effect on the variability and level of output when imposing a zero bound on the short-term interest rate. But output suffers when inflation targets range between 0 and 1%. Clearly the 2% value has great significance as a threshold point in their model. Nevertheless, the authors do not attempt to formally relate the 2% threshold to any underlying macro or microeconomic factor(s).

In this paper, I show that an inflation rate near 2% naturally arises as the product of financial innovations and the pursuit of price stability. In line with the transaction motive literature (Baumol, 1952; Tobin, 1956), I argue that money reduces transaction costs as compared to barter. The novelty is that the transaction cost savings can be directly inputted in terms of the loss of real GDP/capita that would occur if the economy reverted to barter. The fraction of real output saved is assumed to be a function of technical progress and of the cost of substitutes of money, in particular credit cards.

The introduction and expansion of credit cards in the world economy has been a major factor for speeding-up the velocity of money (Geanakoplos and Dubey, 2009). I demonstrate that transaction cost savings vary positively with the net return on assets for depository institutions. This is not trivial and the argument requires a separate and extensive analysis of the interaction between the markets for credit card loans and general “bank” loans. Putting these elements together, and given that the transactional cost savings of money are reduced when more money-like substitutes (credit cards) are used, I find that there must be a positive relation between transactional cost savings and the financial sector’s net asset returns.

_4_ This is related to the concept of net interest margin. The difference is that the return on net assets uses equity as the base, whereas the net interest margin uses total assets.
On the other hand, technical progress has also helped reduce the transaction costs associated with barter, “primitive” forms of money\(^5\) and fiat money as well. For example, the creation of a fully electronic banking clearinghouse system such as ACH (1994) has speeded-up interbank settlements, and the use of electronic money has boomed with online shopping and banking, due to the development of web browsers (1994).

I use real GDP per-capita (labor productivity) as the variable representing technical progress. I distinguish between two main categories of progress: \textit{regime-biased} and \textit{regime-neutral} progress. Regime-biased technological progress enhances the relative efficiency of a given form of money by comparison to barter, which leads to ever greater savings from using that form of money. A regime-neutral technological innovation renders \textit{all} forms of money and barter more efficient, and therefore the relative savings do not change.

I derive the optimal aggregate quantity of money per-capita (and velocity of money) by equating the marginal value of cost savings to the opportunity cost of holding money \textit{relative} to barter. I find that the optimal velocity of money is declining with the net return on assets for depository institutions and thus is also a declining function of a parameter representing the long-term leverage of these institutions. Furthermore, the log of the velocity of money is a linear function of the log of real GDP/capita. If technical progress in the transaction technology is slightly regime-biased towards the current form of money, the model implies that the velocity of money rises at a rate close but less than long-run real GDP/capita growth (2.19\% over the period 1959-2007).

I conduct empirical tests of these relations using Johansen’s (1988, 1991 and 1995) VECM to estimate the long-run equilibrium for the U.S. velocity of narrow money over the period 1959-2007. Since the mid-1980s, and due to changes in the behavior and components of this aggregate, M1 is no longer considered the appropriate measure of narrow money. I test the model using alternate narrow money measures (M1RS and M1S) developed by Dutkowsky, Cynamon and Jones (2006). Overall, the results lend strong support to the model. In all cases, I find that progress indeed has been new regime biased, and that the growth rate of velocity is estimated at 1.91\%.

\(^5\) The idea has a well established tradition. Richard Cantillon, John Locke, Knut Wicksell, Irving Fisher, and Milton Friedman all pointed to innovations as a factor speeding-up the velocity of money (Humphrey, 1993).
Notwithstanding, as M1 does not escape reserve requirements, the connection between M1 velocity and depository institutions’ *actual* leverage appears clearer and more stable than when M1RS and M1S are used. This is because, in the aggregate, (credit card) loans must be a multiple of demand deposits, where the multiple is closely related to the actual (maximized) leverage used by depository institutions. On the other hand, when taking into account the deposits that escape reserve requirements for example due to sweep programs, I recover a value for the estimated leverage parameter that is close to imposed limits historically set by the Fed, the FDIC and international agreements on capital adequacy ratios (Basel I and II), when using the new aggregates M1RS and M1S.

I then show that if the money supply follows a “naïve” Friedman (1960) k-percent rule that aims at achieving price stability, the long-run rate inflation is also a simple linear function of the long-run real GDP/capita growth. In other words, given that the velocity of money expands due to innovations, when monetary authorities pursue a M1 growth target that matches the growth in real GDP, then a long-term inflation near 2% is implied. Based on the velocity cointegrating equations, my point estimate for the long-run inflation rate equals the growth rate of velocity; again 1.91%. Contrary to the widespread belief that money growth rules are incompatible with interest rate targeting rules, this particular k-percent rule is fully consistent with a derived Taylor (1993) type rule.

The rest of the paper proceeds as follows. In section 2, I develop the transaction cost savings function and justify its form in the context of the historical evolution of money. In section 3, I develop the optimality condition leading to the determination of the optimal velocity as well as the optimal quantity of narrow money. I introduce the VECM framework for the empirical tests in section 4. Section 5 features the tests for various measures of narrow money (M1 and adjusted measures of M1). Section 6 demonstrates how the optimal long-run inflation rate is tied to real GDP/capita growth. Section 7 shows the compatibility of a version of Friedman’s (1960) k-percent rule with a new Taylor-type rule. Section 8 discusses Friedman’s (1969) deflationary monetary policy proposal in the context of the new interest targeting rule derived in section 7. Possible extensions are discussed in the concluding section.

---

6 I term this a ‘naïve’ k-percent rule because monetary authorities have historically treated the velocity of money as a random walk with zero drift (Bordo and Jonung, 1987).
2. The Evolution of Money and Transaction Cost Savings

Without a standard of exchange, transaction costs would rise steeply in the economy. A pure barter economy constrains trade and production. James Tobin (1992, p. 18) writes “Does an economy arrive at the same real outcomes... as it would with the institution of money? Clearly not. Without money, confined to barter, the economy would produce a different menu of products, less of most things. People would spend more time searching for trades and less in actual production, consumption and leisure.”

In The Age of Turbulence, Alan Greenspan (2007, p. 2) reflects that: “We’d always thought that if you wanted to cripple the U.S. economy, you’d take out the payment system... Businesses would resort to barter and IOUs; the level of economic activity across the country could drop like a rock.”

In general, transactions costs associated with barter originate from 1) bookkeeping and coordination costs due to tracking relative prices for great variety of goods and services (menu costs); 2) the cost of storing perishable goods for the sole purpose of trading, which leads to waste, and 3) the search cost associated with finding bilateral trades or achieving double-coincidence of wants. The thesis of this paper is that each one of these costs manifests as a reduction of real GDP/capita from its otherwise ‘potential’ level achievable with the institution of money.

In a commodity-money (CM) economy, money is a physical good produced in limited supply or supplied at a slow rate because of high marginal costs of prospection and extraction. While technical progress can push down per-unit costs, it is clear that society must divert resources away from consumption and investment goods in order to produce the commodity money solely for transaction purposes (Friedman, 1951). Of course, this cost must be smaller than the cost associated with a barter economy, otherwise society would not adopt CM. Friedman estimates the cost of production of commodity money at about 1.5% of GDP. In what is possibly the clearest thinking on the topic of commodity-money standards, Friedman (1951, p. 210) writes: “… in a world in which total output is growing in response to technological and other changes and in

---

7 Tobin (1992) is careful to distinguish between the real effects due to the absence of a standard medium of exchange vs. the use of a numeraire. Of course, changing the units of account (new numeraire) has no real effect. This is the classical neutrality of money proposition.
which the velocity of circulation is fairly constant, a strict commodity standard requires the regular use of a considerable volume of resources for additions to the monetary stock in order to keep prices stable.” And in the adjacent footnote he adds “…The naïve notion that there is an immediate and direct “economy” in the use of gold achieved by making gold a fraction of the total circulating medium is obviously wrong.”

Historically, barter and CM economies have experienced flat or even negative real per-capita economic growth. Based on GDP and population data compiled by Maddison (World Tables, 2001) the average nominal GDP/capita growth was a miserly 0.02% per year over the period 1 AD to 1700, about 50 years prior to the first industrial revolution. During that span of time, inflationary periods alternated with deflationary periods. Historians have documented that Western Europe experienced a deep deflation in the second quarter of the 14th Century, then a drastic inflation in the three decades that followed (Munro, 1984). In England, the price surge was further fueled by a series of coinage debasements. No other severe deflation occurred, at least in England, over the period covering the 1400s until the 1870s.

When real GDP/capita growth is flat or even negative, transactions are not increasing in volume so rapidly as to outstrip the supply of commodity monies, which leads to stable prices. Thus, CM continues being used as long as the marginal intrinsic value of one unit of CM is greater than its marginal cost of production. With the industrial revolution and the corresponding boost to productivity growth, governments found it more difficult to maintain ease of trade with metal monies. This is because the production of the commodity money was not keeping pace with economic growth. They faced the prospect of secular deflation and the ensuing risk of economic contractions, as long as money was either an actual commodity or was backed by one (Friedman, 1951; Bordo and Filardo, 2005).8

---

8 Economists have recently revisited the conventional wisdom regarding the economic impact of deflations. Bordo and Filardo (2005) distinguish between what they term good, bad and ugly deflations. They classify the period from 1921-1929 as a good deflation with a mild price decline of 1-2% per year. Good deflations are initiated by boosts in productivity growth. However, it is not clear whether even good deflations may start that way and end-up being bad deflations because of a reduction in aggregate demand due to hoarding money and nominal cuts in wages leading to prices spiraling down and economic slowdown. The policy reaction of monetary authorities obviously matter a lot. The 1928 Fed policy tightening is viewed as responsible for turning the good deflation into the economic contraction of 1929-1933.
Of course, an increase in the velocity of money could have taken care of this problem. But there are technological constraints to increasing the velocity in a pure CM economy. The banking clearinghouse system experiences bottlenecks when the volume of transactions multiplies rapidly due to the resource cost of CM production, the cost of transporting physical money and the limited capacity of commercial banks to use the money multiplier (Norman, Shaw and Speight, 2006). Technical progress did not provide much relief to reduce these costs either.

The logical solution was to introduce fiat money (FM). A transition phase in which paper money was backed by precious metal reserves was the natural course taken by many economies. This stage was a natural extension of banking via the issuance of goldsmiths’ certificates. It was also important for economic agents to gain trust and acceptance for this new form of money. However, that stage could have been bypassed because it eventually leads to the same problem as a strict commodity standard.

For example, during the Gold standard era with backed paper money, the deflationary period of 1873-1896 appears to have been caused by a drop in the growth of gold production while the production of other goods accelerated. The great gold rush in U.S. and Australia of the 1850s had run its course by the 1860s. From 1870 to 1896, the general level of prices declined at an average rate of 1.1% per year. Deflation arose as a consequence of too few money chasing too many goods, because the stock of paper money was in fixed proportion to gold (and silver) reserves.

From its creation in 1913, the Federal Reserve System has served as an increasingly efficient clearinghouse for the banking sector and has taken the role of supplier of liquidity of last resort. This has insured that interbank settlements are no longer the major

---

9 Even though they were able to create inside money, in the form of IOUs to partially circumvent that problem.

10 Short-term deflationary episodes could also happen as result of speculation or arbitrage. For example, the passage of the Sherman Act in 1890 effectively put the United States on a bimetallic currency standard. When the value of silver started dropping relative to gold, economic agents exchanged their silver for dollars (at a legal fixed rate) and in turn demanded gold for their dollars. This riskless arbitrage continued to drain gold from the U.S. monetary system and provoked a financial panic in 1893. On the other hand, temporary inflation can also happen for several reasons: either as a result of central banks temporarily reneging on their commitment to convertibility and/or printing too much money, or as a result of an increase in supply due to discoveries of new Gold deposits and innovations in mining technology. Inflationary periods have in general been short-lived although innovations in gold mining technologies and new mines discoveries did have longer lasting inflationary effects. For instance, major gold discoveries in the Klondike, South Africa, and Australia provided a large boost to money supply after 1896. As a consequence, the U.S. general level of prices increased at 2.5% per year over the period 1896-1914.
source of potential frictions for the velocity of money, which is not to say that credit 
crunches cannot cause a stalling of the velocity of money as experienced during the worst 
episode of the 2007-2010 financial crisis in the fall of 2008. The next big step has been 
the invention of the mainframe computer followed by that of credit cards. Both 
inventions ushered a new era of rapid innovation in banks’ back office operations, 
financial instruments and transaction methods as well as interbank settlements (see Table 
2).

To sum-up, the costs savings associated with using money relative to barter are 
derived from:

1. Economizing on the resource costs due to the lack of double coincidence of wants. 
   This occurs in a barter economy. Resources are wasted due to search costs and 
   storage/spoilage costs.

2. Economizing on the resource cost to produce commodity money. In a barter 
economy, this cost is incurred repeatedly to sustain most of the stock of “ mediums of 
exchanges” needed for transactions, as many of these barter goods are perishable. In a 
commodity (metallic) standard, production costs only occur for new money, and 
constitute a deadweight loss for society in the sense that resources are transferred to a 
non-productive and non-consumable goods sector of the economy. In the transition 
from barter to CM, the cost savings are getting smaller over time because the 
marginal cost of producing CM is rising. In a FM economy the cost of producing 
money is essentially zero.

3. Economizing on the costs generated by an inefficient banking clearinghouse system. 
   In a commodity money system which maintains convertibility, clearing transactions 
   with physical money settlements is costly. The fractional reserve system of the past 
   and our current FM system have drastically reduced these costs.

4. Avoiding recessionary deflations: deflations caused by secular money supply 
   constraints have typically being associated with economic slow downs. Guerrero and 
Parker (2006) find that a higher rate of deflation reduces the subsequent economic 
growth rate (even if it does not always lead to recession). Thus, there is reason to 
believe that deflation is bad for economic growth even if it has become a relatively 
rare experience for most developed economies in the postwar era.
5. Advances in transaction methods: such as credit cards and online shopping and banking, which boost the velocity of money.

It seems clear that the step-wise evolution of exchange systems from barter to CM, and from CM to FM has provided incremental cost savings. Again, the key idea here is that the cost savings of using money and associated transactional innovations can be directly inputted in terms of avoiding loss of real GDP. Given a prevailing monetary system (CM or FM), going back to a lesser form of money or barter would lead to a permanent drop in real GDP. Let me now develop the notations for the paper:

- \( M_t \) = Stock of nominal money
- \( y_t \) = Real per-capita GDP
- \( V_t \) = Velocity of money
- \( y_t \) = Nominal GDP
- \( P_t \) = GDP deflator
- \( N_t \) = Population
- \( m_t \) = Real money per-capita

\( M_t \) are aggregate nominal balances carried from the beginning of period \( t \) (time \( t-1 \)) enabling the purchase of goods and services in period \( t \). The variable \( m_t = \frac{M_t}{P_t N_t} \) is the real money holdings per-capita. In this case, I consider that the relevant head count (i.e. per-capita) includes the population of individuals, businesses and non-federal government entities holding accounts at depository institutions.\(^{11}\) I thus define the average society member as a representative domestic economic agent who is a composite of all these categories.

**Definition (Transactional Cost Savings of Using Money):** Given a current level of real per capita output \( y_t \) achieved by the economy, let \( y_t^B \) denote the level of real GDP/capita (in a barter economy) that would prevail if the institution of money (but not credit) was abolished; i.e. the use of the real (per-capita) stock of money \( m_t \) were suddenly suspended in period \( t \). I denote by \( T_t = y_t - y_t^B = A_t \times y_t \) the real cost savings of using \( m_t \) and associated transactional innovations, where \( 0 \leq A_t \leq 1 \) is the fraction of current real GDP/capita that would be lost if the economy reverted to barter.

---

\(^{11}\) I ignore the issue of demand for currency due to trade balance/current account.
The function $T_i$ captures the current state of the transaction technology. The larger the value of $T_i$, the greater the advances in the transaction technology as compared to barter, and the greater the savings are by comparison to barter. This definition is flexible enough to account for various monetary regimes such as CM and FM.

**Definition (Unit Costs Savings Function $A_t$):** The per-unit-of-real-goods cost savings function is given by

$$A_t = A \times (1 + R_r^A)^{1-\lambda} \times (1 + R_r^F)^{1-\lambda} \times I_t^{1-\lambda_1} \times m_t^{1-\lambda_4}$$

with $0 \leq \lambda_3, \lambda_4 \leq 1$.

The function $A_t$ accounts for three key features of a transaction technology. A transaction technology is characterized by: 1) the medium of exchange used in the economy, i.e. the monetary (barter) regime; 2) the availability of substitutes for the medium(s) of exchange, and 3) the type of technological progress affecting one’s ability to transact for a given amount of real money per-capita. Feature 1) is addressed by using the size of real money per-capita in circulation. In that case, zero money holdings means that barter is in place. I assume that the function $A_t$ has decreasing marginal returns in the amount of real money per-capita, which is equivalent to assuming that the per-unit-of-money transaction (or average) cost savings function $\frac{T_i}{m_i}$ is decreasing with the amount of real money per-capita.

To model feature 2) above, the transaction cost savings function $A_t$ is made to depend on the net cost of using money substitutes (e.g. inside money in the case of barter and commodity money; or credit cards in FM economy). Given my focus on the post-1950s era in the empirical section, I only consider the effect of credit cards loans on transaction cost savings. The transaction cost savings of holding money (per unit of...

---

12 I define a transaction technology as the medium of exchange (money or barter goods) and the associated devices or methods used to facilitate transactions (for example credit cards are associated with fiat money).

13 Overnight interbank lending and central bank loans to depository institutions are also factors that can enhance the velocity of money by facilitating interbank settlements. The demand of overnight funds to meet interbank settlements corresponds to unforeseen imbalances. It appears to be highly interest inelastic over the “normal” range of discount and federal funds rates. In the U.S., the supply of these short-term loans is essentially “perfectly” elastic due to the commitment of the Fed to meet the short-term reserve/liquidity needs (not excess reserves needs) of depository institutions at the prevailing discount rate, especially when the federal funds market does not. Thus, the Fed partly fulfills its role as lender of last resort. The equilibrium quantity of short-term loans thus can potentially adjust to any level required by demand shocks. A reference on the topic is Freixas et al. (2009). As the Fed has historically imposed administrative rationing, institutions prefer using the federal funds market during the normal course of...
goods) \( A_t \) must be decreasing with the availability of more credit instruments that closely substitute for money.\(^\text{14}\)

In Appendix A, I analyze the interaction between the credit card market and the “bank” credit market. I conduct a comparative statics exercise and show that as credit becomes more expensive, the demand for credit cards actually expands while the net return on depository institutions’ assets declines. The reason for the expansion of credit cards is that consumers and businesses view the offered “high” rates on credit cards as price ceilings. They believe they can control the effective interest they pay on the loan over the grace period by avoiding finance charges (Ausubel, 1991). Thus, while other sources of credit are becoming more expensive, that is not necessarily the case for credit cards if customers are disciplined enough to avoid finance charges. Putting these elements together, I show that there must be a positive relation between the cost savings \( A_t \) and the financial sector’s net asset returns, holding real money balances constant.

I assume that the log of the cost savings function, or percentage deviation of cost savings from trend, varies positively with depository financial institutions’ net return on assets. The net return on assets is captured by the geometric compounded average return of the long-term interest rate \( R_t^L \) and short-term interest rate \( R_t^S \), and where the exponents represent the respective weights of assets (loans) and liabilities (deposits) in the financial institutions’ balance sheet.

\footnote{There is no reason why credit cards could not exist and function properly in a pure barter economy. But of course the advent of credit cards has raised the velocity of fiat money. Assuming that money holdings do not change, this leads to an increase of aggregate demand and real GDP/capita, at least in the short-run. To get back to Tobin’s (1992) point though, new transactional technologies have enabled large amounts and variety of goods to be transacted. Thus, it is hard to dissociate the use of credit cards with the achievements of our modern economies in terms of high levels of real GDP/capita. In other words, any transaction technology should be viewed as an integral part of the aggregate production function. Hence, while the per-unit of goods transaction cost savings is decreasing, the total transaction cost savings function should be increasing in the availability of credit cards as they allow larger levels of real GDP to be sustained. Holding the amount of real money per-capita constant in the transaction cost savings function is possible in the presence of credit expansion when credit cards are used as a pure substitute to money holdings. However, as Stauffer (2003) points out, the majority of card users carry a positive balance on average, which boosts M1. Also, it is important to recognize that as the net asset return rises due for example to a reduction in short-term rates, the quantity of other credit products (e.g. mortgages) will rise and substitute for credit card loans within the financial sectors’ loan portfolio.}
In general, depository institutions borrow short-term and leverage-up to lend at long-term interest rates. The institutions’ net return on assets thus can be proxied by 
\[
\left[(1-\lambda_1)R_L^S - \lambda_2 R^L_L\right].
\]
In this context, the parameter \(\lambda_1 = -\lambda_2 > 1\) naturally represents the long-term leverage of depository institutions; i.e. the ratio of total assets over equity. For example, a value \(\lambda_1 = 2\) means that institutions borrow another 100% of their equity at short-term rates to double up the return on long-term loans.\(^{15}\)

The final feature of a transaction cost savings function is addressed by introducing the variable \(I_t\), which represents an index of technical progress. At first glance, the unit cost savings function \(A_t\) being decreasing with the index of progress is counterintuitive. However, recall that both the index of progress and real GDP/capita are variables in the total cost savings function \(T_t\). A surge of technological progress while holding real GDP/capita constant implies that the use of other factors such as labor input must be shrinking. In that case, the cost savings relative to barter are actually diminishing for people, as they have more free time on their hand and better technology, which reduces the cost of bilateral searches, for example by using faster computers to engage in barter. Notwithstanding, this scenario does not really arise in this model as I assume that technical progress is labor augmenting. Hence, progress tracks real GDP/capita (labor productivity) or \(I_t = y_t\).\(^{16}\)

Whether the cost savings function \(T_t\) is increasing in real GDP/capita then depends on the net effect of real GDP per-capita growth via the income and the technical progress

\(^{15}\) While it is beyond the scope of this paper, an analysis of inside money substitutes; i.e. IOUs used to settle interbank payments in a pure CM economy can be accommodated by my transaction cost function. In that case, the portfolio weights should be equal to one another in absolute value; i.e. \((1-\lambda_1) = \lambda_2 > 0\), and thus the transaction cost savings is a decreasing function of \(R^S_L - R^S\); i.e. the lower the profit from lending to non-bank customers, the more willing banks are willing to lend to each others. This is expected in a pure CM economy, with no credit product that can easily substitute for money. Overall, the intuition behind my results is similar to Wicksell’s (1936) analysis of money velocity that revolves around the development of credit as a substitute for commodity money. In his case however, money is exclusively seen as metallic currency or gold. Wicksell was later abundantly criticized for treating bank notes and bank deposits as different from money.

\(^{16}\) I recognize that along a steady-state growth path for economy, it is essentially impossible to separate out the economic effects of real GDP per-capita growth vs. that of labor-augmenting technical progress on transaction costs savings, as in the long-run, these two variables are inextricably tied. My analysis does not focus on transitional dynamics, but on steady-state growth. In that case the sole engine of economic growth is technical progress.
channels. This in turn depends on the sign of the parameter $\lambda_3$. I interpret the parameter $\lambda_3$ as an indicator of the type of technical progress in the transaction technology. I distinguish between *regime-biased* or *regime-neutral* innovations. Regime-biased technical progress reduces transaction costs associated with a specific monetary regime.

Biased progress can be of two subtypes: 1) *New-regime biased* $\lambda_3 > 0$: technological progress enhances the efficiency of the current monetary regime more than barter; 2) *Old-regime biased* $\lambda_3 < 0$: barter receives a greater efficiency boost than the current form of money, or alternatively, technological progress cannot prevent rising costs of producing money. Regime-neutral progress ($\lambda_3 = 0$) allows all transactional systems (from barter to the current form of money) to benefit equally from a new technology. In other words, the efficiency gap is constant. This is obviously a knife-edge case. Table 2 summarizes some of the highlights in the history of major innovations that transformed transaction technologies over the 1959-2007 period. There, I also offer a rough classification of each innovation according to whether it is neutral or biased.

3. **The Optimal Quantity and Velocity of Money**

In a classic paper, Friedman (1969) demonstrates that the socially optimal amount of money to hold is at the point where the marginal benefits of holding money equal its opportunity cost (nominal interest rate) plus the marginal cost of producing money. In a fiat money economy, the marginal cost of money production is zero. Friedman assumes that economic agents have a satiation point for holding money, which means that the marginal gain becomes zero beyond a certain threshold. Consequently, the only way the

---

17 It is possible that more primitive types of money also receive an efficiency boost. Assuming $A > 0$ rules out the case where the old type of money leapfrogs the newer type.

18 Note that the cost savings function must satisfy $0 \leq A_t \leq 1$. Using the Quantity Theory of Money equation expressed as $M_t \times V_t = Y_t$, the real stock of money per-capita is given by $m_t = y_t / V_t$. Thus, the function $A_t$ can also be written as $A_t = A \times (1 + R_t^d)^{-\lambda_3} \times (1 + R_t^s)^{\lambda_3} \times V_t^{\lambda_3-1} \times m_t^{\lambda_3-\lambda_4}$. As long as the velocity of money is not falling to zero and the stock of real money per-capita is not declining, the condition $\lambda_3 < \lambda_4$ plus the proper bound on the constant $A$ are sufficient to insure that $0 \leq A_t \leq 1$. In the limit case where money velocity drops to zero, this constraint can still be met as long as real money per-capita grows faster than the speed of velocity decline.
marginal gain of holding money can equate its marginal cost is by having monetary authorities set the short-term interest rate to zero.\textsuperscript{19}

Even though Friedman (1969) does enumerate the advantages of money over barter, he does not explicitly incorporate these benefits in his analysis. Presumably, these benefits do not have to be analyzed separately as they are already included in the pecuniary services of money. However, Friedman implicitly assumes that a full-blown monetary system is \textit{already} institutionalized. The average household not holding money does not imply that money is completely absent from the economy as a medium of exchange, and that this household has to resort to barter. In fact, economic agents can recover money by selling goods or less liquid assets.

Here, on the other hand, I quantify the marginal costs and benefits of using money \textit{relative} to barter. Money must fulfill the same basic transactional services as barter in addition to removing the frictions caused by barter. Indeed, there are sizable transactional frictions associated with holding goods for barter rather than using money as discussed in the previous section.\textsuperscript{20} The frictions associated with barter are possibly removable not by using a special tax scheme or setting the short-term interest to zero, but rather by achieving a sufficiently high level of technological progress.

To set-up the first result, it is also important to stress that there is a \textit{relative} opportunity cost of holding money as compared to holding goods for barter. The credit/lending technology is \textit{less} efficient in a barter economy due to storage and spoilage costs. Given a rate of interest $R$ promised in a monetary economy, if the monetary system suddenly collapses, the expected return in a barter economy becomes $(1+R)(1−δ)−1$, where $0 ≤ δ < 1$ is the per-dollars storage and spoilage cost.\textsuperscript{21} Thus, the \textit{relative} opportunity cost of holding an amount of money (or goods equivalent to the amount of

\textsuperscript{19} Hetzel (2007) points out from the title Friedman (1969) chose for the section of the book on the optimal rule, that Friedman did not appear to intend for this rule to serve as an actual guide for monetary policy.

\textsuperscript{20} Along Friedman’s line, Wolman (1997) focuses on individual money holdings. He considers the marginal savings in terms of wage earnings due to less time spent transacting. Lucas (2000) also focuses part of his paper on developing a model that incorporates a shopping time constraint, again within a monetized economy.

\textsuperscript{21} I assume that spoilage costs affect interest paid and principal in a pure barter economy, so that $(1+R)(1−δ)−1≥0$. This is similar to comparing (in the standard macroeconomic growth model) the interest rate in an economy where capital depreciates vs. the case of zero depreciation. With zero depreciation, $(1+R)$ equals the marginal productivity of capital. With positive depreciation the new rate is $(1+R_{\text{new}}) = (1+R−δ)$. However it becomes $(1+R_{\text{new}}) = (1+R)(1−δ)$, if the interest earned (in goods terms) is subjected to spoilage as well, for example due to having to store principal and interest before payment.
money) for trade is therefore higher in a monetized economy than in a barter economy by the amount \( \delta(1+R) \).

**Proposition 1**: Assume that there exists a representative unit basket of goods and a quantity \( B \) of that basket satisfying trades in a barter economy. Assume that the consumption in autarky \( C^A \) is related to the quantity of baskets bartered as follows \( C^A = [V^B - 1]B \); where the velocity \( V^B \geq 1 \) applies to a monetized economy at the point of collapse, i.e. experiencing a level of real output per-capita equal to \( y^B \). Further assume that the parameters of the transaction cost savings function satisfy \( \lambda_3 < \lambda_4 \). The optimal quantity of money per-capita and optimal velocity for an economy producing \( y \) are given by:

\[
\ln(m_t) = \frac{1}{\lambda_4} \ln \left[ \frac{A \times (1 - \lambda_4)}{\delta} \right] + \frac{\lambda_4}{\lambda_4} (R^L_t - R^S_t) + \frac{\lambda_3}{\lambda_4} \ln(y_t) \quad (1)
\]

\[
\ln(V_t) = \frac{1}{\lambda_4} \ln \left[ \frac{\delta}{A \times (1 - \lambda_4)} \right] - \frac{\lambda_4}{\lambda_4} (R^L_t - R^S_t) + (1 - \frac{\lambda_3}{\lambda_4}) \ln(y_t) \quad (2)
\]

**Proof**: Assuming that the opportunity cost of money is given by the short-term interest rate \( R^S_t \), the standard optimality condition determining the optimal money holdings \( m_t^* \) is \( \text{Net Marginal Benefits of Real Money Holdings} - R^S_t = 0 \). On the other hand, the services that money renders must at least be equal to the transaction services obtained by holding goods for barter, as money is an extension of barter. In other words, money services can be separated into two additive components: 1) the same basic transactional services that goods held for barter provide, plus 2) the reduction of transactional frictions caused by barter.\(^{22}\)

The general optimum quantity of money condition above can be re-written by segregating the marginal benefits due to barter and those incremental savings due to instituting a monetized economy at the point of collapse: \( \text{Net Marginal Benefit of Holding Barter Goods} - (1+R^S_t)(1-\delta)+1 + [\text{Net Marginal Benefit of Real Money Holdings} - \text{Net Marginal Benefit of Holding Barter Goods}] - \delta (1+R^S_t) = 0 \); where the

\(^{22}\) Here, I assume that the marginal benefits of holding barter goods are net of current spoilage costs and that the transaction cost savings from using money are net of the cost of producing money.
bracketed term corresponds to the marginal cost savings associated with real money holdings relative to barter, i.e. \( \frac{\partial T}{\partial m} \).

Assume that the institution of money collapses and the economy suddenly reverts back to barter. The quantity of money equation must hold true at the point where the economy transitions into barter.\(^{23}\) People must be indifferent between holding a given amount of real money \( m_i \) and bartering and thus \( m_i V_i^B = y_i^B = C_i^A + B_i \); with \( V_i^B \geq 1 \) representing the velocity of money in that economy. Here, the average real output is the sum of consumption in autarky \( C_i^A \) and the representative basket of goods held for barter \( B_i \) per-capita, which includes goods purchased by consumers and producers. At the point of collapse, people must be able to purchase the total consumption, which splits into i) what they can eventually produce and consume in autarky plus ii) additional goods they are looking to barter for.

I assume that the consumption in autarky is related to the amount of goods bartered in the following way \( C_i^A = \left[ V_i^B - 1 \right] B_i \). This condition is an equilibrium resource constraint. Each real unit of money must be ultimately redeemed against its real good (barter) equivalent only once. Of course, the last person left holding the (money) bag loses, if money is viewed as worthless. As Ritter (1995) points out, there are two possible ways out of this conundrum. First, the government maintains the convertibility of money into goods/commodities, for example by raising tax revenues in kind (i.e. non-perishable commodities). The other, is that if convertibility is not maintained, the government commits to not use seignorage, so that money can be stored away and used again sometimes in the future. Another reason why money still has value, is that autarky production has not started yet, so that people do not have the necessary spectrum of goods to barter effectively. For example, State and federal employees or software engineers and teachers would probably have a hard time bartering the services they produce, and so some economic activities would naturally be impeded. The raison d’etre

\(^{23}\) A case in which money and barter can co-exist for a while is that of hyperinflation. A good reference for a model of the transition from barter to fiat money is Ritter (1995). Ritter’s model is based on Kyatoki and Wright (1993).
for the velocity $V_t^s$ being greater than one here is to accommodate this purchase of “autarky” goods before autarky production begins. This latter condition combined with the quantity equation implies $m_t = B_t$. 24

In a pure barter economy, the net marginal benefit of holding one additional unit of real barter goods held as inventory is the value of the incremental goods to be purchased, which are not available in autarky. At the optimum, this value must be equal to the opportunity cost (interest rate net of spoilage due to storage) or $(1 + R_t^S)(1 - \delta) - 1$. The representative basket with $B$ units of barter goods provides a utility of $U(B)$ due to welfare improving trades. The optimality condition that determines the optimal holdings of barter goods $B^*_t$ is

$$\frac{\partial U(B_t)}{\partial B_t} \bigg|_{B^*_t} = \left(1 + R_t^S\right)(1 - \delta) - 1.$$

The optimality of the quantity of unitary basket $B^*_t$ must be satisfied independently from the monetized economy. Because during the sudden transition to barter, the condition $m_t = B_t$ holds, the optimal real money balances must also satisfy the same optimality condition with respect to providing equivalent transaction services to that offered by the basket of real barter goods. Because money holdings $m^*_t$ were optimized before the collapse of the fiat money economy, this leads to $m^*_t = B^*_t$. Hence, the optimum quantity of money $m^*_t$ can be determined simply by equating the marginal benefits vs. the opportunity cost of holding money in excess of those generated by barter; that is:

$$\frac{\partial T_t}{\partial m_t} = A \times (1 - \lambda_t) \times (1 + R_t^S)^{1 - \lambda_t} \times (1 + R_t^L)^{1 - \lambda_t} \times y_t \times m_t^{1 - \lambda_t} = \delta(1 + R_t^S).$$

Rearranging the terms and taking the logs leads to expressing real money balances as

$$(1).$$

From the QTM equation, the real stock of money per-capita is given by $m_t = \frac{y_t}{V_t}$.

Using this last equation and taking the logs, the velocity of money can finally be

---

24 My approach implicitly necessitates the presence of a medium of account in the barter economy (unit basket of goods), which can translate into the real money unit. For example, the medium of account could be a mixture of the reference baskets in the CPI and PPI indexes, and the unit of account is 1 unit of that mixed basket. For a discussion of the difference between medium of exchange vs. medium of account vs. unit of account, see McCallum (2003).
expressed as (2) above. The assumption $\lambda_3 < \lambda_4$ guarantees that the velocity of money rises with real GDP/capita, our index of progress. QED.

The optimal velocity of money given in equation (2) is increasing in per-capita income due to technical progress and decreasing in the return differential $(R^L_t - R^S_t)$. Recall that the spread $(R^L_t - R^S_t)$ shrinking is equivalent to the net return to assets dropping while the quantity offered of credit cards is actually increasing. The optimal velocity being decreasing with $(R^L_t - R^S_t)$ makes economic sense because the velocity of money rises with the expansion of credit cards in the economy.

The optimal quantity of money given by equation (1) is increasing in per-capita income and decreasing in the short-term interest rate as one should expect. On the other hand, it is increasing in the long-term interest rate. Again, this effect is due to the fact that holding the amount of money per-capita constant, depository institutions are expanding the velocity of money via supplying more credit cards loans. The novelty is the inclusion of two new key parameters in the optimum quantity of money equation. These are the degree of bias in technical progress and the leverage ratio of depository institutions. Adrian and Shin (2009) for example, make a strong case in favor of including leverage as a variable in monetary policy rules. Here, leverage is constant as my focus is on a long-run equilibrium.

Equation (1) is consistent with the semi-log money demand function used by Bailey (1956) and Friedman (1969), by contrast with the double-log schedule preferred by Lucas (2000). Here, the per-capita real income elasticity of money demand is expected to be positive and close to zero, due to the fact that progress in the transaction technology is new-regime biased. The optimum quantity of money is a function of the long-term and short-term interest rates, which is unusual in the literature, but defended by Brunner and Metzler (1989).

Contrary to several recent models replicating Friedman’s (1969) optimum quantity of money result, I do not assume satiation in this model in order to avoid the unrealistic
prediction of infinite money holdings at zero interest rates.\textsuperscript{25} This is simply a standard feature of the semi-log schedule I use here, that a finite level of cash is held at a zero nominal interest. The (short-term) interest elasticity of money demand is \[ \left[ -\lambda_i / \lambda_i \times R^S_i \right], \] which is decreasing as the interest rate falls to zero. This simple result is in agreement with Mulligan and Sala-i-Martin’s (1996) findings. They use a cross-sectional data on asset holdings of U.S. households in 1989 and find that the interest elasticity is very low when interest rates are low. They state on page 41: “Our prediction of low interest elasticity at low interest rates is crucial, for example, for the evaluation of the welfare costs of inflation. The consumer surplus approach applied by Bailey (1956), Lucas (1994) and others show that the welfare cost of inflation hinges fundamentally on the money demand elasticity at low interest rates.” Marty (1999) also criticizes Lucas’ choice on the basis that it artificially inflates welfare gains since the level of cash balances goes to infinity as the interest rate approaches zero.

My model is most related to the literature on innovation and money. Ireland (1994) argues that the effects of economic growth on the payments system can be substantial. In his model, the ratio M2/M1 rises steadily over time and the demand for M1 becomes increasingly interest-elastic as the economy grows. The fact that M2/M1 rises is fully consistent with my hypothesis that substitutes of narrow money enhance the velocity of narrow money. Ireland (1995) models the process of financial innovations as carrying a fixed cost. In order to innovate, the opportunity cost of holding cash balances has to be higher than a threshold level that might be greater than current interest rates. The model is based on Dotsey’s (1984) framework in which financial innovations are endogenous and treated as investment projects. In a different vein, Jafarey and Masters (2003) study the impact of innovation on the velocity of money within the framework of a search/matching model. When the matching technology is improved they show that the velocity of money is affected positively.

\textsuperscript{25} Wolman (1997) assumes satiation in money holdings in order to reproduce Friedman’s (1969) result. Mulligan and Sala-i-Martin (1997) assert that the assumption of satiation can be done away with in their model as long as seignorage falls as the interest rate drops to zero. In other words, the stock of money does not grow faster than the speed of decline of the interest rate. Some cash-in-advance models appear to predict Friedman’s result without resorting to assuming satiation. However, they assume that frictionless barter is achievable as a first best solution. Recently, a new string of the literature integrates market frictions in hybrid models combining general equilibrium and search models (Aruoba et al., 2007). These models reach conclusions at odds with Friedman’s (1969) recommendation.
4. Data and Testing Methodology

In this section, I test the long-term optimal velocity of money relation (2) for the U.S. on a quarterly basis over the period Jan. 1959- Oct. 2007. The variables used are real GDP per-capita calculated using real GDP (GDPC96; 2005 dollars, seasonally adjusted annual rate) from BEA and U.S. population is monthly data (POP) from the Census Bureau that is matched with real GDP observations on a quarterly basis. Long-term interest rates are monthly rates based on the long-term bonds series (LTGOVTBD) from the Fed’s BOG, which is the unweighted average yield on all outstanding bonds neither due nor callable in less than 10 years. The series stops in June 2000. I complete the long-term bond series by using the constant maturity (monthly) 10-year Treasury yield (GS10) from Q3 2000 to Q4 2007. I use three alternate money stock measures related to M1. The first measure is M1 from the Fed’s BOG (M1SL), which is monthly and seasonally adjusted data. The other two measures are M1RS and M1S available on a seasonally adjusted basis only up until the end of 2007. This data is from Cynamon, Dutkowsky and Jones’s website at http://www.sweepmeasures.com/.

Since the 1970s depository institutions have been able to lower their reserve requirements on commercial demand deposits by sweeping deposits into other instruments, but initially the size of these operations was never great due to the lack of computer speed, given that the sweeps had to be returned at regular intervals. Since 1994, depository institutions have been able to lower their reserve requirements for retail deposits by using automated computer programs to move inventories of checkable deposits overnight into money market depository accounts (MMDAs) and money market mutual funds (MMMFs), which are not subject to reserve requirements.

These operations known as sweep account programs have led to a situation where M1 is underestimating actual narrow money. Anderson (2003) points out that sweep programs are initiated by banks not by account depositors. Depositors optimize their cash holdings with the understanding that the whole balances are available from the bank at any point in time for transaction needs. Thus, it makes sense to include sweeps in a measure of narrow money.
Recently, Anderson (1997) and Dutkowsky and al. (2003) estimated the magnitude of these sweep programs, and in particular Dutkowsky et al. (2006) developed M1RS and M1S, the two new adjusted measures for M1. The new aggregate M1RS equals M1 + (swept funds from retail programs), which are funds with unrestricted transaction properties. M1S equals M1RS + (swept funds from commercial demand deposits sweep programs), which contains all sweeps.\(^{26}\)

By definition, the actual velocity of money is \(V_t = \frac{Y_t}{M_t}\). Figure 1 shows the velocities of all three measures of narrow money and Figure 2 shows the velocity of M1 in relation to real GDP/capita growth. Visually, Figure 2 illustrates that there seems to be a strong connection between the log of M1 velocity and the log of real GDP per-capita growth over the period. On the other hand, from Figure 1, such a connection appears to break down with M1RS and M1S, as these money aggregates expanded rapidly after the early 1990s.

I test the long-run optimal velocity equation (2) using the following specification for the equilibrium relation:

\[
\ln(V_t) = \ln(\alpha) + \beta_1 \times R_s^t + \beta_2 \times R_d^t + \beta_3 \times \ln(y_t) + \beta_4 \times d597t1 + \beta_5 \times SM1_t, \tag{4}
\]

My goal is to estimate the key coefficients \(\ln(\alpha), \beta_1, \beta_2, \) and \(\beta_3\). These coefficients are related to the optimal velocity equation (2) in the following way:

\[
\ln(\alpha) = \frac{1}{\lambda_4} \ln\left(\frac{\delta}{A \times (1 - \lambda_4)}\right); \quad \beta_1 = \frac{\lambda_4}{\lambda_4} > 0; \quad \beta_2 = -\frac{\lambda_3}{\lambda_4} < 0 \quad \text{and} \quad \beta_3 = 1 - \frac{\lambda_4}{\lambda_4} > 0. \quad \text{I employ Johansen’s (1988, 1991 and 1995) Vector Error Correction Model (VECM) approach. The reason is that the variables may have a unit root and are possibly cointegrated. First, I confirm that all the variables are I(1).}^{27} \quad \text{I also test the log of velocity series for structural}
\]

\(^{26}\) Cynamon, Dutkowsky and Jones’s database tracks commercial demand deposit programs only after 1991. Thus, M1S may be slightly biased downward prior to 1991.

\(^{27}\) I run Augmented Dickey Fuller tests with trends for the log of velocity of money (M1, M1RS and M1S) as well as for SM1 and the log of real GDP per-capita, and with drift for the two interest rates. In all cases the null hypothesis of non-stationarity cannot be rejected at the 1%, 5% and 10% levels, except for the T-Bill rate, which is rejected at the 5% level but not at the 1% level. The detailed results are obtainable from the author. The other possible issue is the collinearity between the two interest rates used here. However, rather than imposing restrictions, I let the VECM structure speak for itself and determine whether there is an independent cointegrating equation governing the behavior of these two rates, with short-term adjustments away from equilibrium. The empirical results do not support that contention. Any short term
breaks. Using Clemente et al. (1998) I test for the presence of unit roots with at most two breaks. I find that indeed two breaks are present in the log of velocity of M1RS and M1S. Based on whether the test is for additive outliers (AO) or innovational outlier (IO), I find that for M1RS the break dates are (Q2 1976; Q2 1994) for AO and (Q3 1971; Q3 1991) for IO. In the case of M1S, the break dates are (Q2 1976; Q2 1994) in the AO case, and (Q3-1971; Q3 1991) in the IO case. Interestingly, one of the two extreme break dates Q3 1971 represents the end of the gold standard for the U.S. economy and the other Q2 1994 is one quarter after retail sweep programs began. I choose Q3 1971 as the earliest break date in the analysis for both M1RS and M1S. For M1RS I choose the beginning of the retail sweep program or Q1-1994. For M1S I choose Q3 1991 for the second break, that is, when total sweeps started being recorded.

The idea behind Johansen’s (1988, 1991, 1995) method is to estimate the full dynamic structure of the relationship between these variables, while at the same time being able to separate out the long-run from the short-run dynamics. In this paper, I am only interested in studying the long-run dynamics part of the VECM. The set up is

$$\Delta x_t = \theta(\beta x_{t-1} + \mu) + \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} + \gamma + w_t d5971 + w_t d7982 + \varepsilon_t.$$  

Where the vector $x_t = (\ln(V_t), R^S_t, R^L_t, \ln(y_t), d5971, SM1_t)'$ is (6x1), $\theta$ is a (6xr) matrix representing the speed of adjustment, and $\beta$ is a (rx6) matrix representing the parameters of the cointegrating equations $(\beta x_{t-1} + \mu)$. The number of cointegrating relations is $r$. The matrices $\Gamma_i$ are (6x6) and $p$ is the number of lags in the short-term dynamics. I restrict the cointegrating equation to be stationary around constant means $\mu$, which is a (rx1) vector. I also allow for a stochastic trend in levels by including a constant (6x1) vector $\gamma$ in the short-run dynamics part of the VECM.

The (6x1) vector $w_t$ is associated with the dummy variable $d5971$, which accounts for a possible break in the times series of narrow money velocity due to the transition from the gold standard to pure fiat money economy after the Nixon administration ended the dollars-gold peg. It also accounts for the reduction in the volatility of real GDP/capita relationship between the two rates will be uncovered by the short-term dynamics of the VECM and thus separated from the cointegrating relation (4).
post gold standard era. This dummy variable takes a value of 1 from Q1 1959 to Q3 1971 and 0 otherwise.

The variable $SM1$ is the ratio of sweeps divided by $M1$. This variable is used in the VECM only when the velocity of money is calculated using $M1S$ or $M1RS$, for the purpose of accounting for structural breaks in these series. The $SM1$ variable spans January 1994 to October 2007 for $M1RS$ and October 1991 to October 2007 for $M1S$. The $(6 \times 1)$ vector $w_2$ is associated with a seasonal dummy variable $d7982$, which accounts for period covering the high inflation of 1979 and the 1982 recession under Volcker’s tenure. This dummy is only assumed to affect the short-term dynamics of the VECM. For the sake of simplicity, I assume for now that the number of cointegrating relations is $r = 1$, which I verify later. Using Johansen’s (1995) normalization procedure, so that the coefficient on $\ln(V_i)$ is normalized to 1, the matrix $\beta$ can be expressed as $\beta = (1, -\beta_1, -\beta_2, -\beta_3, -\beta_4, -\beta_5)$ and $\mu = -\ln(\alpha)$, which leads to the same form as equation (4) above.

5. Test Results with M1 and Other Measures of Narrow Money

Table 3 presents the results for several versions of the cointegrating equation (4) presented above. I test various lag structures as well. Each specific version presented in Table 3 minimizes either the AIC (Akaike criterion), the BIC (Bayesian information criterion) or the HQ (Hannan-Quinn information criterion) across possible lags from 1 to 4. All equations included in Table 3 have a cointegration rank of one; i.e. only one cointegrating relation exists in each case. All coefficients for the long-run equilibrium are significant at least at the 99% level, except in a couple of instances. Overall, the basic model is supported for each measure of narrow money used here.

For M1, the best result supporting the model is when the dummy variable $d5971$ is present in the model and the VECM lag = 2. In that case, the AIC, BIC and Chi$^2$ statistics

---

28 The definition is different for $M1RS$ and $M1S$. For $M1RS$, the variable $SM1$ only includes retail sweeps. For $M1S$ it includes total sweeps.

29 I excluded instances of rank equal to 2, which occurred mostly when the number of lags equaled 4. These multiple cointegrating equations were economically meaningless. I also ran post estimation diagnostics on the VECM. Overall, the presented equations were part of VECMs which had stable roots, with normally distributed disturbances (except for the short term dynamics of the log of real GDP/capita), and exhibited some residual autocorrelation at all lags up to lag 4. The results are available from the author upon request.
are minimized. The estimate for the elasticity of real income per-capita is 0.85 and the estimate of the leverage coefficient is 14.91 (for $R^S$) and -15.09 (for $R^L$). These two coefficients are nearly equal in magnitude as predicted by the theory.

Interestingly, the value of these coefficients is close to the actual historical leverage of depository institutions. Figure 3 graphs the actual leverage for depository institutions over the period 1959-2007. Since the late 1980s, the leverage has drastically dropped. This corresponds to the adoption of the several rounds of Basel accords regarding capital adequacy ratios, starting in 1988. The mean of the time series equals 15.01, which is close to the leverage coefficients found above. However, recall that the estimated parameter is the ratio of leverage $\lambda$ over the parameter $\beta$ so that $\beta = \frac{\lambda}{\lambda_4}$. In this case, there is reason to believe that the coefficient on the long-term rate is more exact, as it does not suffer from the distortion created by the introduction of NOW account in the early 1980s. The introduction of these accounts made the decision to hold cash less sensitive to short-term interest rates. Thus, I use $\beta = 15.09$ and $\lambda = 15.01$. The coefficient $\lambda_4$ can be estimated at 0.99. This indicates that the elasticity of real money $(1-\lambda_4)$ in the transaction cost savings function is very close to zero.

The results with M1RS and M1S are not as strongly supportive but they are still consistent with my findings for M1. In the case of M1RS the best result in support of the theory is when the VECM lag = 4 and the dummy $d5971$ is used in conjunction with the variable $SM1$, i.e. retail sweeps divided by M1. In that case, the long-run relationship shows an elasticity of income equal to 0.84. The values for leverage coefficient are respectively 12.64 (for $R^S$) and -13.19 (for $R^L$). The two coefficients are still close to each other in absolute value. Equally good with respect to the AIC criterion is the model that includes the dummy $d7982$. In that case the coefficient is 11.43 (for $R^S$), -12.82 (for $R^L$) and 0.91 for the income elasticity. The best result for M1S, which minimizes the AIC

---

30 I consider that leverage ratios with a value above 100 are outliers, and therefore are removed from the sample. These are banks and financial institutions with the following NAICS codes: 522110 commercial banks, 522120 savings institutions, 522130 credit unions and 522190 other depository credit intermediation institutions. After 1999, I add other institutions due to the 1999 banking deregulation. The additional codes are: 522210, 220, 291, 292, 294, 390; 523220, 524126, 127 and 210.

31 I discuss later in this section why M1 is a better choice to capture the actual leverage parameter.
criterion is with a lag = 4 and including the \(d5971\) dummy and the ratio of total sweeps over M1. In that case, the income elasticity is 0.89 and the two leverage coefficients are respectively 13.70 (for \(R^5\)) and -13.90 (for \(R^7\)).

Taking the average of the four estimates for each measure of narrow money, my point estimate for the coefficient \(1 - \lambda_3 / \lambda_4 = 0.8725\), so that \(\lambda_3 = 0.12\) and technical progress in the transaction technology is slightly new-regime biased. By the same token, I derive an average estimate for the constant \(A\) in the per-unit transaction cost savings function that equals 1.1. This is assuming that the depreciation (spoilage) rate in a barter economy has a value around 0.45% per year which is on the low side as compared to a rate of 1% per year estimated from grain agriculture in today’s rural China (Park, 2006), which is the only estimate I found in the literature. Under this assumption, the per-unit transaction cost savings function \(A_t\) is less than 1 over the sample period, as required by definition.

The reason I include the variable \(SM1\) in the model is to account for the break in the two series after 1991 and 1994, respectively when commercial demand deposits and retail sweep programs were instituted. Algebraically it is easy to show that:

\[
\ln(VRS) = \ln(V) - \frac{\text{Retail Sweeps}}{M1} - \ln(V) - SM1.
\]

That is, the log of the velocity of M1RS equals the log of the velocity of M1 minus \(SM1\). The same identity holds when M1RS is replaced by M1S. In the case of M1RS with a lag = 4, the coefficient on the \(SM1\) variable is not equal to 1 but rather it is \(-0.69\). This means that the variable \(SM1\) is having a distorting impact on other variables in the cointegration equation. An increase in the \(SM1\) ratio correlates with a lower income elasticity and a lower leverage than would be generated by M1 alone.

Even though computer software executes tasks faster and more efficiently, depository institutions still have to return swept funds back to the account they were swept from at set intervals, due to regulations and customer cash needs. Thus, as institutions optimize to keep the sweeps as large as they can, the value of sweeps cannot outpace the value of deposits in the long-run, and thus sweeps must become a constant fraction of narrow money. Figure 4 visually shows that the ratio of retail sweeps to aggregate M1RS has

---

32 I run VECM “control” experiments without dummy variables and as can be seen from the bottom half of Table 3, none of the experiments conform to the predictions of the theory and there is no discernable pattern for any of the money measures.
been leveling off over time. As the ratio converges to a constant, sweep programs will only have an intercept effect to lower the velocity of money. Sweep programs should then have no “pervasive” effect on the income elasticity, but may still impact the leverage parameter at the margin. In Figure 4, however, total sweeps as a percentage of M1S do not exhibit that leveling-off pattern… yet!

I find that the velocity elasticity with respect to the short-term interest variable is smaller than the elasticity with respect to long-term interest, for M1, M1RS and M1S. One explanation is that since NOW accounts were instituted after 1981. These are checking-type accounts that pay interest, with some restrictions however, as for-profit corporations are excluded from opening these accounts. As a result, the (short-term) opportunity cost of holding cash balances is lower than it was before.

As sweep programs distort the relationship between required reserves ratio and demand deposits, it turns out that reverting to M1 as the measure of narrow money does actually help to uncover a stable long-run relationship between banks’ actual leverage ratios and the velocity of money, as all M1 funds are subject to the Fed statutory reserve requirements and to the capital adequacy ratio rules from the 1988 Basel and later accords.33

Because institutions are able to shelter some of their assets from regulatory requirements, they are effectively able to leverage themselves at a higher level than if all assets in M1RS or M1S were counted. This means that the leverage ratio associated with M1 should be closer to actual ratios, and the estimates associated with M1RS and M1S should be lower as compared to M1. Furthermore, the Basel accords are enforced over the entire asset base of financial institutions, whether they use sweeps or not. Given that assets are risk-weighted to compute the required leverage ratios, the minimum of 8% (equity/assets) will lead to an effective leverage ratio that is slightly higher than 12.5 (inverse of 8%) but also lower than the historical average of 15.01, which is what I find for the leverage estimates using M1RS and M1S.

33 Interestingly, as more lenient reserve requirement rules were enacted since the 1990s, one should have expected leverage to increase in depository institutions, but in fact one observes the opposite in Figure 3. This might be due to the counteracting effect of the Basel Accords regarding capital adequacy, especially the dip after 1990.

A recent wave of articles examines the new measures of narrow money M1RS and M1S developed by Dutkowsky et al. (2006). Dutkowsky et al. test the existence of a long-run demand relationship over 1959-2002 using M1, M1RS and M1S. They do not include a measure of progress in their equation. They assume that the income elasticity of demand is unitary, which means that the velocity only depends on the rate of interest, a standard assumption in the literature. Not surprisingly, their best result is achieved when using M1S, which has the least amount of trending amongst the three alternate measures of narrow money. Ireland (2008) extends the work of Lucas (2000) and uses M1RS to measure the welfare cost of inflation. His conclusion weighs in favor of a semi-log formulation for the optimum quantity of money.

6. A Constant Money Growth Rule and a Near 2% Inflation Target

To derive the optimal long-run inflation rate, I combine my derived optimal velocity of money with a Friedman’s (1960) k-percent rule for monetary policy. I assume that monetary authorities use a naïve Friedman k-percent rule in the sense that velocity is treated as a random walk with zero drift (Bordo and Jonung, 1987) so that money supply is set to grow at the same pace as real GDP capita in the long-run.

Proposition 2: In a steady-state, assuming that a naïve Friedman k-percent rule applies; that is \( g_M = g_y + n \) and that transactional innovations are new-regime biased on average \( (\lambda_3 \geq 0) \), then the long-term rate of inflation is \textit{constant} and equals the per-capita income elasticity times the long-term real GDP/capita growth \( (1 - \frac{\lambda_3}{\lambda_4}) \times g_y \).
Proof: The QTM equation in terms of long-term growth rates is 
\[(1 + g_m) \times (1 + g_y) = (1 + \pi_x) \times (1 + g_y) \times (1 + n).\] 
Assuming a ‘naïve’ Friedman k-percent rule is equivalent to 
\[(1 + g_m) = (1 + g_y) \times (1 + n).\] 
Both equations together imply that 
\[g_y = \pi_x.\] 
As the short-term and long-term nominal interest rates are constant in the steady-state, the optimal velocity equation (2) implies that 
\[g_y = (1 - \lambda_3 / \lambda_4) \times g_y.\] 
Combining with the previous relationship this implies that 
\[\pi_x = (1 - \lambda_3 / \lambda_4) \times g_y.\] 
QED.

The long run real GDP/capita growth rate is estimated at 2.19% over the period 1959-2007. Using the point estimate from Section 5 above I find that the optimal long-term rate of inflation has a value of \[0.8725 \times 2.19\% = 1.91\%,\] which matches the rate of growth of money velocity. Next, I explore the feasibility of a link between a long-term money growth objective and interest rate targeting.

7. A Long-Term Taylor-Type Rule Compatible with a Money Growth Rule

It is my view in this section that a Taylor rule is effectively the central bank’s reaction function to optimal money holdings. In other words, assuming that the Fed knows the optimal quantity of money function, it can indirectly impact it by acting on the interest rate. If the central bank’s long-term objective is price stability, it is crucial that it understands well the actual behavior of the long-term optimum quantity of money function. Here, I derive a Taylor-type rule using the optimal quantity of money equation (1) rewritten as follows: 
\[\ln(m_t) = C + \lambda_1 / \lambda_4 \times (R_t^L - R_t^S) + \lambda_3 / \lambda_4 \times \ln(y_t).\] 
Where the constant 
\[C = 1 / \lambda_4 \times \log \left[ A \times (1 - \lambda_4) / \delta \right].\] 
I assume that there is a long term optimal amount of money \(\bar{m}_t\) that corresponds to an equilibrium value for both the short-term and the long-term interest rates \(R_t^S\) and \(R_t^L\) and the potential real GDP/capita \(\bar{y}_t\). Therefore, the optimum is defined by 
\[\ln(\bar{m}_t) = C + \lambda_1 / \lambda_4 \times (R_t^L - R_t^S) + \lambda_3 / \lambda_4 \times \ln(\bar{y}_t).\] 
Assume monetary authorities set out to pursue price stability by following a naïve k-percent Friedman type rule where nominal money growth equals long-run real GDP growth. That rule is equivalent to minimizing the difference between the log of the supply of real
money per-capita and the long-term optimal level, or setting \( \ln(m_t) - \ln(\bar{m}_t) = 0 \), which leads to

\[
R^S_t = \bar{R}^S + (R^L_t - \bar{R}^L_t) + \frac{\lambda_5}{\lambda_1} \times \left[ \ln(y_t) - \ln(\bar{y}_t) \right].
\]

To recover values close to Taylor’s (1993) rule, I assume along with Taylor (1998) that the long-term rate satisfies the Fisher effect, then \( \bar{R}^L = 2\% \) (real interest) + \( \bar{\pi} \) (inflation target), where \( \bar{\pi} = 1.91\% \). Furthermore, I assume that the current long-term rate is \( R^L_t = 2\% \) (real interest) + \( \pi_t \), where \( \pi_t \) stands for the expected inflation rate applicable to the current long-term nominal rate. In that case, I find that

\[
R^S_t = \bar{R}^S + (\pi_t - \bar{\pi}) + \frac{\lambda_5}{\lambda_1} \times \left[ \ln(y_t) - \ln(\bar{y}_t) \right].
\]

Equation (5) is the key result of this section. The coefficients of this Taylor-type rule have economic meanings, which is not typically the case in the literature. In this context, the key coefficient of equation (5) is \( \frac{\lambda_5}{\lambda_1} \), which is the ratio of the technological bias parameter divided by the leverage ratio of depository institutions. The naïve k-percent rule in equilibrium implies that the short-term interest rate must be set to the natural rate \( \bar{R}^S = \bar{R}^L \), which in Taylor’s case is set at 2\% real (assuming a 2\% inflation). Taylor’s (1993) coefficients on the inflation and the output gaps are both 0.5. In my case, the coefficients are 1 for the inflation gap, and 0.8\% for the log of the output gap. Of course, the comparison is not apples to apples as Taylor analyzes only a 5-year period, and I use a 50-year period. He uses the federal funds rate as the short-term rate, whereas I use the 3 month T-Bill. Moreover, he uses real GDP not of real GDP/capita as I do here. The coefficient on the output gap is quite small in my case, which means a fairly insensitive response of policy to the business cycle.

Because the two rules are essentially equivalent, why would the Fed not directly implement a money growth rule? In hindsight, the breakdown of money growth rules after 1982, pointed to the fact that the transmission channel did not account for shifts in velocity, and financial innovations. Possibly going the route of a Taylor rule may be easier if the parameters of the optimal velocity function are stable enough.

It is important to emphasize however that the Taylor-type rule I derived assumes that the velocity of money is in its long-run equilibrium. This rule may be appropriate for example when the intent is to smooth interest rates and policy adjustments are gradual.
(Dueker, 1999). It is not necessarily appropriate for sharp short-term policy responses to exogenous shocks, such as driving short-term rates to zero to avoid a recession. Notwithstanding, contrary to previously held beliefs, I am able to reconcile a ‘naïve’ k-percent Friedman type of rule with a Taylor-type rule. Orphanides (2007) states: “A policy rule quite as simple as Friedman's k-percent rule cannot be formulated with an interest rate instrument. As early as Wicksell's (1898) monumental treatise on Interest and Prices, it was recognized that attempting to peg the short-term nominal interest rate at a fixed value does not constitute a stable policy rule. (Indeed, this was one reason why Friedman, 1968, and others expressed a preference for rules with money as the policy instrument.) Wicksell argued that the central bank should aim to maintain price stability, which in theory could be achieved if the interest rate were always equal to the economy's natural rate of interest, r*.”

By contrast, I find that a k-percent money growth rule and a Taylor-type rule are interchangeable. Nevertheless, caution must be exercised when applying the naïve k-percent rule in its interest targeting form. Setting the real short-term interest to its long-term counterpart would constrain the real term structure to be flat. This might be a problem as this works against the segment of investors who are trying to hedge short-term risk, and are willing to bid up short-term Treasuries and accept a lower real yield than that of long-term instruments. Faugere and Van Erlach (2009) for example show that since the mid-1950s real after-tax short-term one-year Treasury yields have embedded a negative time-varying risk premium by comparison to 30-year Treasuries.

8. A Near 2% Target to Avoid “Bad” Deflations

Whether or not the policy shift to a Taylor-type rule was done with that purpose in mind, a Taylor rule has the advantage of being more flexible by contrast with money supply growth objective, as it leads to a money supply that adjusts to short-term variations in the velocity of money and even to the case where velocity growth may stall in the long-term. Bordo and Filardo (2005) for example state that “When inflation is low, the usefulness of

---

34 An exception is Taylor (1998) who uses the QTM to informally derive a general Taylor type rule based on a constant money growth rule. But he does not specify nor does he derive a money demand function.
35 There is a literature contrasting the effects of either rule in terms of their impact on economic stability (Evans and Honkapohja, 2003).
monetary aggregates may be exceeded by that of short-term interest rates, especially if velocity is sufficiently unpredictable.”

Whereas Bordo and Filardo may be more concerned with short-term policy, avoiding long-term recessionary deflations is important as well. Assuming that the pace of innovation in the transaction technology slows down and velocity becomes constant (but not zero) then the Taylor type rule defined in Section 7 generates price stability, i.e. an actual inflation rate of 0%. Hence, with the help of this policy, actual inflation is mostly contained between 0% and 2%. On the other hand, velocity may temporarily accelerate faster than the 1.91% pace, as it did in the mid 1990s (M1 velocity). In that case, interest targeting leads to slightly greater inflation than desired. But this is a temporary situation because the pace of financial innovation must revert back to long-term productivity growth.

A different but related literature has examined what inflation level is needed to avoid a recessionary deflation in the context of monetary policy hitting the zero nominal bound. Coenen et al. (2003) build a model for a small open economy with staggered wages subject to stochastic shocks similar in magnitude to those experienced in the U.S. over the 1980s and 1990s. Once shocks to aggregate demand or supply push the economy into a sufficiently deep deflation, a zero-interest-rate policy may not be able to return the economy to the original equilibrium. With a series of shocks large enough to sustain deflationary expectations and to keep the real interest rate above its equilibrium level, aggregate demand is suppressed further sending the economy into a deflationary spiral. They find that the consequences of the zero bound are negligible for target inflation rates as low as 2 percent but not lower. By contrast, my analysis clearly shows that interest targeting provides a long-term hedge against deflation, whether the zero bound is present or not.

On the other hand, Friedman (1969) argues that a rate of deflation equal to the negative of the real interest rate might be desirable. Over the past 20 years, Friedman’s (1969) proposal has certainly been extensively analyzed by a plethora of macro models.36

36 A classic article that predicts Friedman’s result is Abel (1987). Mulligan and Salai-i-Martin (1997) provide a great survey in which they find that Friedman’s conclusion is model dependent. Of course, theoretically Friedman’s (1969) proposal is subject to the same limitations as mentioned before; given the assumptions of 1) satiation and 2) pre-existing monetized economy. Also it should be logically clear that
While, this prescription has been clearly rejected by major central banks as a guide for conducting long-term monetary policy, some low level deflation may be acceptable and even desirable, as long as it is accompanied by productivity increases and no severe downward spiral in nominal wages and aggregate demand.

As the thesis of this paper argues, financial and technical innovations matter in determining the behavior of money velocity. While this is not a new idea, it seems that practically speaking, this point has been ignored by policymakers. The implication of my analysis is that a slow down of the velocity of money may accentuate the rate of deflation, turning it into a “bad” deflation as economic agents join the vicious cycle of money hoarding followed by economic slowdown. I am not here pinpointing the particular threshold where this would happen. However, it appears that the potential risks associated with Friedman’s (1969) recommendation are great and that central bankers have used their judgment correctly in staying away from it.

9. Conclusion and Extensions

In this paper, I provide a macroeconomic foundation for why the velocity of M1 behaved as it did over the past 50 years, as well as for why an inflation target near 2% is optimal. Even though policymakers may not be able to implement a precise target in practice, defining a credible and appropriate inflation target can help insure that economic expectations are firmly anchored, which for example has strong implications for the stability of assets prices.37

I model the transaction costs savings of money relative to barter as a function of technical progress and of the net return on assets for the depository institutions, which itself depends on leverage. I logically arrive at the conclusion that transaction cost savings depend on the depository institutions leverage ratio due to the expansion of credit cards as a substitute for other sources of consumer/business credit. While the most recognizable feature of credit cards is their ability to substitute for money, I find that the key feature of this transmission mechanism is the grace period, which ultimately drives

---

37 Faugere and Van Erlach (2009) find that the Fisher effect on an after-tax basis is indeed present in equity indexes valuation and Treasuries yields.
the demand by convenience users, and possibly entices credit card users to carry a positive balance due to a lack of vigilance over meeting the card companies’ conditions (Ausubel, 1991).

Using the standard optimality of money holdings condition, I am able to derive the optimal velocity of narrow money, which increases with real GDP/capita and decreases with the net return on depository institutions assets, and leverage. Empirically, I use a VECM approach (Johansen, 1988, 1991 and 1995) and find that for various adjusted measures of narrow money the long run velocity relation features parameter values consistent with the historical record. The leverage parameter in the cointegrating equation (using M1) is a near perfect match with estimates of the mean leverage ratio for depository institutions around a value of 15 over the past 50 years.

As Reynard (2006) remarks, “It is often suggested that an explanation for the upward trend in M1 velocity during the post-war period is that technical progress in credit cards and other advances would have allowed individuals to economize on money balances, justifying an income elasticity below unity.” Here, I do find support for that insight. The income (real GDP/capita) elasticity of velocity is indeed less than unity because progress in the transaction technology is on average biased towards new forms of money.

I assume that monetary authorities use a naïve Friedman k-percent rule in the sense that velocity is treated as a random walk with zero drift (Bordo and Jonung, 1987) and the money supply grows at the same pace as real GDP capita in the long-run. In conjunction with the optimal velocity result, this implies that the optimal rate of inflation is slightly below 2%, and is a function of the bias in progress parameter. My point estimate for the optimal long-run inflation rate is 1.91% for the U.S. Furthermore, I show that a Taylor type rule is fully consistent with a long-run money growth objective.

As it turns out, seeking price stability, under the paradigm that the velocity of money is constant, is not so “naïve” after all. It guarantees that long-run inflation will remain between 0% and 1.91% and thus provides a complete hedge against deflation, would the pace of financial innovations and the velocity of money slow down.

Future research will examine the evolution of money and its impact on velocity as new means of transaction continue to emerge. In particular, given the rising speed of networked computers, there is a resurgence of barter as a method of commerce. Other
possible extensions are to examine the future of non-central bank monies. The impact of securitization is also an aspect that merits further investigation, as the dynamic link between bank profits and the gap between lending rate vs. deposit rate is no longer as obvious, given that financial institutions can partially escape capital adequacy ratios by shipping risky (credit card) loans off their balance sheets.
References


Adrian, Tobias, and Hyun Song Shin, 2009, “Prices and Quantities in the Monetary Policy Transmission Mechanism,” *Federal Reserve Bank of New York Staff Report 396*.


Appendix A: The Credit Card Market and the “Banks” General Credit Market

I show here that the quantity supplied of credit card loans is related to depository institutions’ net return on assets. Notwithstanding, the link is not necessarily as obvious as one might imagine. For example, one might assume that the net return on credit card loans is about equal to that of other loans because of market efficiency, and thus the quantity supplied of credit card loans should rise with the net return on banks assets. However, this is not correct because borrowers substitute between credit instruments to select the “cheapest” one, which turns out to be credit cards, as borrowers have the option of using the grace period to avoid finance charges. This is the core of the analysis done below. Let me first start by covering some of the unique features of the credit card market.

First, it is well documented that during the 1980s and 1990s credit cards profits dramatically outpaced those of other types of bank loans. Ausubel (1997) reports that from 1983 to 1993 the return on assets (ROA) from credit card loans was roughly four times the banks’ overall ROA. On the other hand, data from the Federal Reserve shows that the proportion of revolving loans as a percentage of the total (revolving plus non-revolving) has been rather stable around 37% on average since the early 1990s, with a slight upward trend peaking at about 41% in 1998-1999 and declining since then up until 2010.38 One possible reason why institutions have not expanded the share of credit card loans in their portfolios during that period is that the risk exposure is much greater for these unsecured loans, so that capital adequacy ratios put in place by the Basel accords after 1988 have limited how much of these loans could be placed on the banks’ balance sheets.39

Although it is unresolved at this point whether this is a permanent characteristic of the credit card market or not, Ausubel (1991 and 1997) documents that credit card rates are relatively insensitive to changes in short-term rates (costs of funds). However, they appear to be slightly more sensitive to changes in long-term rates at least since the interest rate ceilings on credit cards were removed in the early 1980s (Brown and Plache, 38 These figures are available from the author upon request.
39 This is assuming the traditional commercial banking model prevails, whereas securitization has changed this situation since the mid 1990s. Currently about 60% of all credit card loans are securitized (Getter, 2008).
Stavins (1996) shows that the demand for credit cards is elastic, so the explanation of interest rate stickiness does not necessarily originate from the demand side, even though in some segments of the market, demand may be more inelastic. For example, the latter is true for customers with high balances, as they face high switching costs (Calem and Mester, 1994). One explanation for the uniformity of “high” rates across the industry is adverse selection. Banks do not want to unilaterally lower their rate, because in doing so they would attract a pool of higher risk customers. On the other hand, the supposedly high interest rates that the industry charges are nowhere near the rate they expect to receive. In 1996, it was estimated that over half and probably as much as 68% of credit card users were considered “convenience users”. These customers use credit cards primarily as a transactional medium and pay off their balances in full each month. Around that same time, Visa estimated that almost 60 percent of total bankcard volume generated no interest. By contrast, revolvers carry a positive balance at the end of the month.

Below I analyze the connection between a drop in the banks’ net asset returns and the equilibrium quantity of credit cards on the market. Assume that the Fed implements a restrictive policy and sets the short-term interest rate at \( R_{0}^{S} > R_{0}^{S} \). Figure 5 describes what happens in the depository institution’s loan market (excluding credit cards). The shift from the supply curve Supply\(_{0}\) to Supply\(_{1}\) takes place as depository institutions provide the same quantities of credit as before at the same maximized profit/net asset return, determined by long-run competition. Here, I assume that the net asset return for non-credit cards loans uses the same leverage ratio than for the total loan portfolio.

Focus on the initial equilibrium \( Q_{0} \). The new interest on loans on the new supply curve \( R_{0}^{L} (R_{i}^{S}) \) must be related to the old interest rate on loans \( R_{0}^{L} (R_{0}^{S}) \) in the following way \( R_{0}^{L} (R_{i}^{S}) > R_{0}^{L} (R_{0}^{S}) \), by the fact that \((1 - \lambda_{i})R_{i}^{S} + \lambda_{i}R_{0}^{S} = (1 - \lambda_{i})R_{0}^{S} + \lambda_{i}R_{0}^{L}, \) where \( \lambda_{i} \) represents the leverage parameter. As long as, the demand is elastic, Figure 5 shows that the new equilibrium rate on loans \( R_{i}^{L} (R_{i}^{S}) \) must satisfy \( R_{i}^{L} (R_{i}^{S}) < R_{0}^{L} (R_{0}^{S}) \) and \((1 - \lambda_{i})R_{i}^{S} + \lambda_{i}R_{0}^{L} < (1 - \lambda_{i})R_{0}^{S} + \lambda_{i}R_{0}^{L}, \) so that \( R_{i}^{L} - R_{i}^{S} < R_{0}^{L} - R_{0}^{S}. \) \(^{40}\) Thus, a contractionary

---

\(^{40}\) The higher lending rates (long-term) will put pressure on the long-term bond market so the yields will go up there too.
monetary policy has the effect of decreasing the net asset return on (non-credit card) bank loans and reducing the quantity of these loans.

The second stage of the argument is to examine what happens in the credit card market. I hypothesize that the market supply and demand of credit cards are “effective” in the sense that they are driven by the weighted average return on these loans, holding the total number of users and the proportion of revolvers vs. convenient users constant.

What is happening here is a substitution effect on the demand side. Because other sources of credit are becoming more expensive, the demand for credit card loans will increase as a substitute, because borrowers believe they can control ex-ante the effective rate they pay. Even if customers have existing balances, they can transfer their balances to new credit cards and benefit from low introductory rates and a new grace period. Ausubel (1991) makes the point that credit card users often underestimate the amount they will borrow, as they are not careful enough to make payments on time, or face unforeseen adverse economic situations. The demand by ex-post convenience users should go up as some of these people will be able to implement the following arbitrage strategy. They can borrow money from the credit cards at a low effective rate (potentially 0%) and lend this capital at a now higher short-term rate. In that case, some individual demands may even have a positive slope. The demand by ex-post revolvers also increases as they end-up with a positive balance, even if they had set-out to have zero balances. Moreover, they can always take out a regular bank loan and pay back their credit card balance, as long as they remain credit worthy. All in all, this results in a displacement of the demand for credit cards.

Figure 6, shows how the demand labeled according to the short term rates \( R_t^5 > R_0^5 \) shifted corresponding to an increased demand mostly driven by an increased proportion of convenience users. The proportion of customers who are “revolvers” is denoted by \( a \). The rotation of the demand inward occurs holding the number of credit card loans constant, after the proportion of revolvers drops from \( a_0 \) to \( a_1 \). On the other hand, a larger number of loans causes the demand \( \text{Demand}_1 (R_t^5) \) to shift outward and end-up where shown in the graph.

Empirically, the maximum rates are insensitive to short-term rates, as discussed above. Assuming stickiness of maximum rates would actually make my argument
I choose to show on the graph an example where the maximum rate charged on credit cards does increase from $R(R_0^S)$ to $R(R_1^S)$. This case corresponds to a shift of an infinitely elastic (probably over a finite range) supply due to the higher cost of funds. This supply corresponds to the “best case” scenario for the banks where all customers would be revolvers. However, that is not what happens in reality. The amount by which the maximum rate shifts is explained below.

The expected/average return on these loans is $E(R_0) = a_0 R(R_0^S)$ before the demand shift and $E(R_1) = a_1 R(R_1^S)$ after. The maximum rates $R(R_0^S)$ and $R(R_1^S)$ charged must satisfy $\lambda_1 a_0 R(R_0^S) + (1-\lambda_1) R_0^S = \lambda_1 a_1 R(R_1^S) + (1-\lambda_1) R_1^S$. In other words, given the leverage $\lambda_1$, when the proportion of revolvers drops from $a_0$ to $a_1$, the maximum rate charged must rise so that the net asset return on credit card loans remains the same as the one determined by the long-run competitive equilibrium. The effective supply is actually shifting up as shown in Figure 6. In that case, banks are not receiving a greater net asset return on credit card loans. In conclusion, a rise in short-term interest rates leads to a substitution of less traditional bank loans in favor of more credit cards loans, at the same time as the net return on banks total assets is declining. A decrease in short-term interest rates leads to the reverse outcome.41

On the other hand, an increase in Treasuries long-term rates first leads to an increase of bank lending rates, because of the two markets competing for funds. Thus, the supplies of bank loans shift (credit cards and non-credit cards). However, because of abnormally large profits, banks bid-up short-term instruments used for leverage. Hence, the supply shift of non-credit card loans is such that the net asset return is again the maximum achievable under long-run competition (holding the quantity of non-credit-card loans constant). Similarly as before, this leads to a decrease in the net asset return (movement along the demand curve) for non-credit card loans. Furthermore, the move upward in short-term rates and long-term rates impacts the maximum card rates. Thus, the credit card market is impacted the same way as before. Finally, I have shown that in all cases,

41 The analysis is done so that the aggregate amount of credit is held constant once the total effects of Figure 5 and 6 are combined. This is a key point because the comparative statics analysis of the transactional cost savings function requires the net asset return to change, holding M1 constant.

45
the equilibrium quantity of credit card loans varies on a one-to-one basis with the net returns on assets for depository institutions.\textsuperscript{42}

\textsuperscript{42} It should be also clear to the reader that the above analysis is consistent with the bank “lending channel” approach to monetary policy as developed by Bernanke and Blinder (1988).
<table>
<thead>
<tr>
<th>Country</th>
<th>Date Adopted</th>
<th>Target</th>
<th>Target variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>1993</td>
<td>Average of 2-3% over the medium term</td>
<td>Underlying PI up until October 1998; CPI thereafter</td>
</tr>
<tr>
<td>Canada</td>
<td>February 1991</td>
<td>Midpoint 2% + 1% band</td>
<td>CPI</td>
</tr>
<tr>
<td>Finland</td>
<td>February 1993</td>
<td>2% no explicit band</td>
<td>CPI excluding indirect taxes, subsidies and housing-related</td>
</tr>
<tr>
<td>New Zealand</td>
<td>April 1988</td>
<td>0-3%</td>
<td>CPI excluding interest</td>
</tr>
<tr>
<td>Spain*</td>
<td>November 1994</td>
<td>2%</td>
<td>CPI</td>
</tr>
<tr>
<td>Sweden</td>
<td>January 1993</td>
<td>Midpoint 2% +1% band</td>
<td>CPI</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>October 1992</td>
<td>2.5% +1% reporting range</td>
<td>Retail price index excluding mortgage interest payments</td>
</tr>
</tbody>
</table>

Table 1
Inflation Targeting by G-7 Central Banks. Source: Reserve Bank of New Zealand.
Table 2: Innovations Applicable to Transaction Technologies

<table>
<thead>
<tr>
<th>Transaction Technology</th>
<th>Sample Period</th>
<th>Fed Chairmanship</th>
<th>Innovations Affecting Money Supply and Transaction Technology</th>
<th>Type of Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>First fully transistorized IBM 7090 mainframe computer (1959)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GE’s ERMA Computer to Process Checks (1959)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>General purpose credit cards (BofA) (1966)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATMs (1967)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Information Management System (IBM, 1968)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Magnetic Swipe Cards (1969)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Microprocessor (1970)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Heap Leach Technology for Gold Mining (1970)</td>
<td>Old Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Clearing House Interbank Payment System (CHIPS, 1970)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Miller 78-79</td>
<td>First PC Altair (1975)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volcker 79-87</td>
<td>Telephone Banking (1979)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Software Standardization of Fedwire System (1980)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Activated Carbon Processes in Gold Mining (1980)</td>
<td>Old Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NOW Accounts (1981), Super NOWs and MMDAs (1982)</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Commodore 64 (1982)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lotus 123 (1982)</td>
<td>Old Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Greenspan 87-06</td>
<td>Windows 1.0 (1985)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>386 chip (1985)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PC-Based Banking Clearinghouse Item Processing System (1986)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Consolidation of Mainframe Computers at Fedwire (1990)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pentium Chip (1993)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bernanke 06-</td>
<td>“All Electronic” Banking Clearinghouse ACH (1994)</td>
<td>Neutral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Online Banking (1995)</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Completion of High-Speed Network FEDNET (1996)</td>
<td>Neutral</td>
</tr>
<tr>
<td>??</td>
<td>Beyond 2007</td>
<td>Bernanke 06-</td>
<td>Digital Money</td>
<td>New Biased</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E-barter</td>
<td>Old Biased</td>
</tr>
</tbody>
</table>
Table 3
Cointegration Results for the Log of Velocity of M1, M1RS and M1S from Jan. 1959 - Oct. 2007 on a Quarterly Basis

The "endogeneized" variable is the Log of money velocity (M1, M1RS or M1S), which following Johansen's method (1995) has a normalized coefficient of 1. All samples start on January 1959 and end October 2007. All trace statistics for the rank of the cointegrating equations reject that the rank is greater than 1 at the 95% level and thus point to a unique cointegrating relation in each row. The Chi² statistics shows that all parameters in the equations are together significant at least at the 99% level. Unless noted, each coefficient in all the cointegrating equations is significant at least at the 99% level. The High and Low columns represent the 95% confidence interval around the estimate of the coefficient for Ln(realGDPc). The AIC, BIC and HQ criteria in the righmost columns indicate the goodness of fit for the VECM.

<table>
<thead>
<tr>
<th>Optimal Lag</th>
<th>Ln Velocity</th>
<th>Trace</th>
<th>R²</th>
<th>Lnreal GDPc</th>
<th>d5971</th>
<th>SM1</th>
<th>Const.</th>
<th>Low</th>
<th>High</th>
<th>Ln Lklhood</th>
<th>Chi²</th>
<th>AIC</th>
<th>BIC</th>
<th>HQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag = 1 SBIC</td>
<td>M1</td>
<td>22.4</td>
<td>16.10</td>
<td>-15.50</td>
<td>0.95</td>
<td>-0.14</td>
<td>---</td>
<td>-1.03</td>
<td>0.76</td>
<td>1.13</td>
<td>2998</td>
<td>448</td>
<td>-30.6</td>
<td>-30.5</td>
</tr>
<tr>
<td>Lag = 2 HQIC</td>
<td>M1</td>
<td>21</td>
<td>14.91</td>
<td>-15.09</td>
<td>0.85</td>
<td>-0.21</td>
<td>---</td>
<td>-0.94</td>
<td>0.64</td>
<td>1.07</td>
<td>3030</td>
<td>265</td>
<td>-30.8</td>
<td>-30.6</td>
</tr>
<tr>
<td>Lag = 4 LR, FPE, AIC</td>
<td>M1</td>
<td>30.2</td>
<td>15.42</td>
<td>-16.92</td>
<td>0.81</td>
<td>-0.30</td>
<td>---</td>
<td>-0.54</td>
<td>0.61</td>
<td>1.00</td>
<td>3049</td>
<td>326</td>
<td>-30.8</td>
<td>-30.6</td>
</tr>
<tr>
<td>Lag = 1 HQIC, SBIC</td>
<td>M1*</td>
<td>21.3</td>
<td>15.21</td>
<td>-15.01</td>
<td>0.93</td>
<td>-0.14</td>
<td>---</td>
<td>-0.92</td>
<td>0.75</td>
<td>1.11</td>
<td>3001</td>
<td>425</td>
<td>-30.6</td>
<td>-30.4</td>
</tr>
<tr>
<td>Lag = 4 LR, FPE, AIC</td>
<td>M1*</td>
<td>27.8</td>
<td>14.05</td>
<td>-15.80</td>
<td>0.80</td>
<td>-0.29</td>
<td>---</td>
<td>-0.48</td>
<td>0.62</td>
<td>0.99</td>
<td>3053</td>
<td>349</td>
<td>-30.8</td>
<td>-30.2</td>
</tr>
<tr>
<td>Lag = 2 HQIC</td>
<td>M1RS</td>
<td>57.9</td>
<td>7.37</td>
<td>-5.06</td>
<td>0.58</td>
<td>-0.15</td>
<td>-0.29</td>
<td>-0.22</td>
<td>0.42</td>
<td>0.74</td>
<td>3810</td>
<td>563</td>
<td>-38.7</td>
<td>-38.4</td>
</tr>
<tr>
<td>Lag = 4 LR, FPE, AIC</td>
<td>M1RS</td>
<td>54.2</td>
<td>12.64</td>
<td>-13.19</td>
<td>0.84</td>
<td>-0.23</td>
<td>-0.69</td>
<td>-0.79</td>
<td>0.58</td>
<td>1.11</td>
<td>3858</td>
<td>267</td>
<td>-38.9</td>
<td>-38</td>
</tr>
<tr>
<td>Lag = 2 HQIC</td>
<td>M1RS*</td>
<td>62.9</td>
<td>7.00</td>
<td>-4.89</td>
<td>0.59</td>
<td>-0.15</td>
<td>-0.30</td>
<td>-0.23</td>
<td>0.42</td>
<td>0.75</td>
<td>3814</td>
<td>267</td>
<td>-38.8</td>
<td>-38.3</td>
</tr>
<tr>
<td>Lag = 4 LR, FPE, AIC</td>
<td>M1RS*</td>
<td>60.7</td>
<td>11.43</td>
<td>-12.82</td>
<td>0.91</td>
<td>-0.22</td>
<td>-0.78</td>
<td>-0.98</td>
<td>0.64</td>
<td>1.19</td>
<td>3864</td>
<td>244</td>
<td>-38.9</td>
<td>-38</td>
</tr>
<tr>
<td>Lag = 2 HQIC, SBIC</td>
<td>M1S</td>
<td>56.1</td>
<td>7.93</td>
<td>-5.40</td>
<td>0.64</td>
<td>-0.13</td>
<td>-0.35</td>
<td>-0.46</td>
<td>0.47</td>
<td>0.81</td>
<td>3765</td>
<td>553</td>
<td>-38.3</td>
<td>-37.9</td>
</tr>
<tr>
<td>Lag = 3 FPE, AIC</td>
<td>M1S</td>
<td>53.6</td>
<td>8.29</td>
<td>-6.42</td>
<td>0.60</td>
<td>-0.18</td>
<td>-0.39</td>
<td>-0.25</td>
<td>0.42</td>
<td>0.78</td>
<td>3794</td>
<td>497</td>
<td>-38.4</td>
<td>-37.8</td>
</tr>
<tr>
<td>Lag = 4 LR</td>
<td>M1S</td>
<td>55.2</td>
<td>13.70</td>
<td>-13.90</td>
<td>0.89</td>
<td>-0.21</td>
<td>-0.58</td>
<td>-0.99</td>
<td>0.61</td>
<td>1.16</td>
<td>3808</td>
<td>275.4</td>
<td>-38.7</td>
<td>-37.5</td>
</tr>
<tr>
<td>Lag = 2 HQIC</td>
<td>M1S*</td>
<td>60.4</td>
<td>7.62</td>
<td>-4.92</td>
<td>0.62</td>
<td>-0.13</td>
<td>-0.33</td>
<td>-0.39</td>
<td>0.45</td>
<td>0.78</td>
<td>3769</td>
<td>470</td>
<td>-38.2</td>
<td>-38.7</td>
</tr>
<tr>
<td>Lag = 4 LR, FPE, AIC</td>
<td>M1S*</td>
<td>59.8</td>
<td>13.00</td>
<td>-13.41</td>
<td>0.88</td>
<td>-0.21</td>
<td>-0.58</td>
<td>-0.92</td>
<td>0.61</td>
<td>1.15</td>
<td>3812</td>
<td>239</td>
<td>-38.3</td>
<td>-38.3</td>
</tr>
</tbody>
</table>

▲ Means that the VECM includes the seasonal dummy d7982
Figure 1: Indexed Log of Narrow Money Velocity (M1, M1RS, and M1S). 1959-2007.

Figure 2: Indexed Log of M1 Velocity vs. Log Real GDP/capita. 1959-2007.
Figure 3: Asset/Equity (leverage) for Depository Institutions. 1959-2007. Source: Compustat. Outlier values above 100 removed.

Mean = 15.01

Figure 4: Retail and Total Sweeps respectively as % of M1RS and M1S 1994-2007
Figure 5: "Bank" Credit Market (excluding Credit Cards)-- Restrictive Monetary Policy

![Graph of Bank Credit Market](image1)

Figure 6: Credit Card Market-- Restrictive Monetary Policy

![Graph of Credit Card Market](image2)