Chapter 2
Money and Credit Demand

A medium of exchange is a good which people acquire neither for their own consumption nor for employment in their own production activities, but with the intention of exchanging it at a later date against those goods which they want to use either for consumption or for production.


2.1 Classical Demand for Money Theory

The quantity theory of money, dating back to contributions made in the mid-16th century by Spanish Scholastic writers of the Salamanca School, is one of the oldest theories in economics (de Soto, 2006, p. 603). In his book The Purchasing Power of Money (1911), Fisher gave the quantity theory, as inherited from his classical and pre-classical predecessors, its modern formulation. Fisher’s version, typically termed equation of exchange or transaction approach can be stated as:

\[ M \cdot V_T = T \cdot P_T, \]  

(2.1)

where \( M \) = stock of money, \( V_T \) = velocity of the stock of money to finance the transaction volume, \( T \), and \( P_T \) = price level.

According to the neo-classical assumptions – namely that the economy is running at full potential and \( V \) is constant – \( P \) would move in strict proportion to changes in \( M \): A rise (decline) in the economy’s stock of money would increase (reduce) the price level. In this theoretical framework, money is neutral as far as its effects on output are concerned. Changes in \( M \) affect \( P \), but do not have any impact on \( Y \) or \( V \).
2.1.1 The Cambridge Approach

The Cambridge approach or cash balance approach is associated with Arthur C. Pigou (1917) and Alfred Marshall (1923). It differs from Fisher’s approach in three aspects. First, the Cambridge approach is a microeconomic approach, describing individual choice rather than market equilibrium. It asks: what determines the amount of money an individual would wish to hold, given that the desire to conduct transactions makes money holding attractive. The Cambridge approach moved the analytical focus from a model where the velocity of money was determined by making payments to one where market agents have a demand for money (Cuthbertson & Barlow, 1991, p. 16). Second, money is held not only as a medium of exchange for making transactions as in Fisher’s case, but also as a store of value, providing satisfaction to its holder by, for instance, adding convenience and security. And third, the concept of money demand comes across more explicitly as will be discussed in more detail below; Cambridge economists pointed out the role of wealth and the interest rate in determining the demand for money.

Formalizing the Cambridge approach, Pigou assumed that for an individual the level of wealth, the volume of transactions and the level of income – at least over short periods – would move in stable portions to one another. Other things being equal, the nominal demand for money, $M_d$, is then proportional to the nominal level of the transaction volume, $PT$:

$$M_d = kPT,$$

where $k$ represents the cash holding coefficient. The latter is simply the reciprocal of the velocity of money, that is: $V = 1/k$. If money supply, $M_s$, equals money demand $M_d$, we can write:

$$M_s = M_d = kPT$$

$$M = \frac{1}{V}PT$$

$$MV = PT,$$

with the latter expression representing the familiar equation of exchange.

The Cambridge formulation of the quantity theory provides a description of monetary equilibrium within the neo-classical model by focusing on peoples’ demand for money, especially the demand for real money balances, as the important factor determining the equilibrium price level consistent with a given quantity of money. The emphasis which the Cambridge formulation placed on the demand for money and implicitly also on its determinants like, for instance, interest rates, strongly influenced both the Keynesian and the Monetarist theories.

1Humphrey (2004) notes that Marshall in his early manuscript Money (1871) as well as in his book Economics of Industry (coauthored with his wife in 1879) gave the quantity theory, as inherited from his classical predecessors, its distinctive Cambridge cash-balance formulation.
2.1 Classical Demand for Money Theory

2.1.2 The Role of Wealth in the Transaction Approach

In most analyses, the velocity of money is calculated on the basis of current output, or gross domestic product (GDP). However, the latter does not take into account the economy’s stock of wealth, or stock of assets (such as, for instance, equities, bonds, housing, etc.). To make explicit the role of the economy’s stock of wealth for the velocity of money, the equation of exchange (2.1) can be written as follows:

\[ M \cdot V_T = \left( Y^\alpha \cdot S^{1-\alpha} \right) \cdot \left( P_Y^\beta \cdot P_S^{1-\beta} \right), \]  

(2.4)

where \( M \) = stock of money, \( V_T \) = velocity of the stock of money to finance the economy’s total transaction volume, \( Y \) = real output, \( S \) = stock of (financial) wealth, while \( P_Y \) and \( P_S \) represent the price level of real output and (financial) assets, respectively. \( \alpha(1-\alpha) \) is the share of real output (financial assets) in the economy’s total transaction volume, and \( \beta \) is the share of output’s (financial assets’) price level in the total price level. Taking logarithms and solving for the velocity of the transaction volume, Eq. (2.4) can be stated as:

\[ \nu_T = \alpha \cdot y + (1 - \alpha) \cdot s + \beta \cdot p_Y + (1 - \beta) \cdot p_S - m. \]  

(2.5)

In most analyses, however, the transaction equation is based on current real output (that is GDP) rather than the economy’s total transaction volume:

\[ M \cdot V_Y = Y \cdot P_Y, \]  

(2.6)

or in log form and solving for the velocity

\[ \nu_Y = y + p_Y - m. \]  

(2.7)

According to this representation, the traditional income velocity of money in log form will be negative (positive) if nominal money supply exceeds (falls below) nominal income. The difference between the velocity of the transaction volume and the velocity of current income is:

\[ \nu_T - \nu_Y = (y - s) \cdot (\alpha - 1) + (p_Y - p_S) \cdot (\beta - 1). \]  

(2.8)

Under the assumption that \( 0 < \alpha, \ \beta < 1 \), which is equivalent to \( \alpha - 1, \ \beta - 1 < 0 \), Eq. (2.8) suggests that the transaction volume velocity of money will exceed the income velocity of money if financial wealth exceeds output \( (y < s) \) and/or the price level of financial wealth exceeds the price level of current output \( (p_Y < p_S) \).

For instance, the income velocity of the euro area money stock M3 – after having shown a more or less stable linear decline since the early 1980s – started falling more strongly as from 2001 (Fig. 2.1). This finding could be an indication that the increase in the stock of M3 increasingly affected nominal wealth (such as, for instance, stock and real estate prices) rather than nominal income.
Fig. 2.1 Euro area income velocity of M3, actual and trend
Source: Thomson Financial, ECB; own calculations. – The income velocity of M3 was calculated by dividing nominal GDP by the stock of M3. Period: 1980-Q1 to 2008-Q1.

Time Series Properties of the Velocity Series

Figure 2.2(a) shows the Japanese income velocity of broad money – defined as the difference between nominal GDP and broad money (in natural logarithms) for the period 1980-Q1 to 2005-Q4. Two linear trend lines for income velocity

Fig. 2.2 Income velocity in Japan
2.1 Classical Demand for Money Theory

are also shown: a trend line for the period 1980-Q1 to 2005-Q4 and a split trend line that runs from 1980-Q1 to 1986-Q4, thereafter adopting a flatter trend decline. Figure 2.2(b) shows the deviations of actual from trend income velocities of money, index, together with the Nikkei stock market.

A visual inspection suggests that the strong rise in the Nikkei around the second half of the 1980s was accompanied by the income velocity of money moving below its long-run trend, while the decline in stock prices in the period thereafter was accompanied by income velocity moving back up towards the long-run trend. That said, the assumed trend figure of the income velocity of money may have important analytical implications.

The time series properties of the income velocity of money affect the analyses related to its medium- to long-term trend. If the log level of income velocity of money, $v_t$, at time $t$ is stationary around a linear trend $t$, i.e.:

$$v_t = a + \beta \cdot t + \varepsilon_t, \quad t = 1, \ldots, T,$$

(2.9)

where $\varepsilon$ is a mean-zero stationary process (a random process which has a constant and time independent variance), then the medium-term velocity developments can be described by estimates of $\beta$ from Eq. (2.9).

If velocity is assumed to be stationary, its variance is constant over time and the covariance between two time instances depends only on the distance, or lag, between them and not on the actual time at which the covariance is calculated. In this case, the impact of shocks on velocity will vanish over time.

In contrast, the log level of velocity is non-stationary if it is a random walk, for instance with drift parameter $\mu$:

$$v_t = v_{t-1} + \mu + \eta_t, \quad t = 1, \ldots, T,$$

(2.10)

where $\eta_t$ is some mean-zero stationary process, and the lagged coefficient on velocity is assumed to be one, which explains why the series is said to contain a unit root.

The impact of a random shock on the velocity of money would never disappear as velocity can equivalently be written as an accumulation of past historical shocks:

$$v_t = v_0 + \mu \cdot t + \sum_{j=0}^{t} \eta_j, \quad t = 1, \ldots, T.$$

(2.11)

As a consequence, the variance of velocity would become a function of time $t$. In theory, then, under the unit root assumption, deviations of velocity from $\mu \cdot t$ would increase over time. However, the size of the random walk
component will also determine how far actual velocity might deviate from the linear trend in a specific period of time.

Under the unit root assumption, velocity can be differenced to obtain a stationary series. Taking the time difference of velocity in Eq. (2.10) and using the specification $\Delta v_t = v_t - v_{t-1}$, yields:

$$\Delta v_t = \mu + \eta_t \quad t = 1, \ldots, T. \quad (2.12)$$

This would permit the estimation of the drift term $\mu$ which could form the basis for the expectation of the money velocity trend over the medium-term.

In sum, the choice between models (2.9) and (2.10) may have important economic implications. In the context of model (2.10), under the unit root assumption, velocity would be treated as a non-stationary time series to which shocks accumulate over time. In model (2.9), velocity would be treated as a stationary time series; shocks would have no permanent impact on velocity. Regressing the level of velocity on a time trend in model (2.9) would be appropriate but not in model (2.10).


### 2.2 Keynesian Money Demand Theory

The Keynesian theory of money demand (or: liquidity preference theory) focuses on the motives that lead people to hold money. More specifically, John Maynard Keynes (1936) distinguished between the demand for transaction balances (including the demand for precaution balances), $L_T$, and the demand for speculative balances, $L_S$:

$$L_T = L_t(Y) = kY, \text{ with } \partial L_T / \partial Y > 0 \quad \text{and} \quad (2.13)$$

$$L_S = L_S(r) = R - dr, \text{ with } \partial L_S / \partial r < 0 \text{ and } \partial^2 L_S / \partial r^2 < 0, \quad (2.14)$$

where $k =$ income balance coefficient, $Y =$ nominal output, $R =$ autonomous speculative balance, $d =$ interest rate elasticity and $r =$ representative interest rate.

Combining the demand for transaction balances (2.13) and speculative balances (2.14) yields the Keynesian demand for money:

---

$^2$“Money held for each of the three purposes forms, nevertheless a single pool, which the holder is under no necessity to segregate into three water-tight compartments; for they need not be sharply divided even in his own mind, and the same sum can be held primarily for one purpose and secondarily for another”. Keynes (1973), p. 195.
2.2 Keynesian Money Demand Theory

\[ L = L_T + L_S = kY + L_S(r). \]  \(2.15\)

In equilibrium, money supply, \(M\), equals money demand, that is, in real terms, \(M/P = L\). Similar to the quantity theory, the transaction demand for money emphasises the role of money as a means of payment, suggesting that the transaction demand for money depends on the level of current income. The store-of-value function is reflected in the speculative motive of the demand for money.

In the Keynesian theory, market agents’ portfolio decisions are driven by expectations regarding future bond prices, e.g. bond yields. Bonds are willingly held if the expected total return (defined as the sum of interest payable on the bond and expected capital gains) is greater than zero.\(^3\)

The pricing formula for the bond market (for consols) is:

\[ BP = \frac{n}{r} \cdot NV, \]  \(2.16\)

where \(BP\) = bond market price, \(n\) = nominal coupon of the bond (in percent of its nominal value), \(r\) = effective yield in the bond market (market rate), and \(NV\) = nominal value of the bond. For example, if \(n = 5\%\) p.a., \(NV = US\$100\) and \(r = 5\%\) p.a., then \(BP = US\$100\). If \(r\) rises to \(10\%\), \(BP\) falls to \(US\$50\).

Now assume that each investor has an estimate of the \(normal\) market yield, \(r_{normal}\), which might deviate from the current market yield. Depending on the subjectively perceived normal yield, an individual investor is expecting a capital gain or loss from bond holdings:

\[ expected \ loss \ on \ bond \ holdings = BP - BP^e, \]  \(2.17\)

with \(BP^e\) as the expected bond price in the future. We can rewrite Eq. (2.17) as follows:

\[ expected \ loss \ on \ bond \ holdings = \frac{n}{r} \cdot NV - \frac{n}{r_{normal}} \cdot NV \]  \(2.18\)

The income from coupon payments on the bond is:

\[ interest \ income = n \cdot NV. \]  \(2.19\)

That said, an investor will hold interest bearing bonds if interest income is higher than expected capital losses from holding bonds:

\[ \frac{n \cdot NV}{interest \ income} > \frac{n}{r} \cdot NV - \frac{n}{r_{normal}} \cdot NV \]  \(2.20\)

---

\(^3\)Note that if a bond’s market price equals its par value, the bond’s return equals its nominal coupon. If the bond’s market price is higher (lower) than its par value, the bond’s return is lower (higher) than its nominal coupon.
or, equivalently:

\[
1 > \frac{1}{r - r_{\text{normal}}} \quad \text{or} \quad \frac{r_{\text{normal}}}{1 + r_{\text{normal}}}
\]

The level of \( r \), which fulfills Eq. (2.20b), is called the \textit{critical market yield}. At the critical yield, the investor is \textit{indifferent} as interest income from bond holdings equals expected capital losses of bond holdings. That said, one can draw the following conclusions:

- If the market rate is \textit{above} the critical yield level, interest income is higher than the expected capital losses, so that the investor decides to keep his total wealth \textit{in bonds}.
- If the market yield is \textit{below} the critical market yield, the investor would decide to keep his total wealth \textit{in money balances}, as he fears capital losses related to a forthcoming rise in market yields.

The Keynesian demand for speculative balances is therefore a decision to hold \textit{either} money \textit{or} interest bearing bonds (Fig. 2.3a). Assuming that different people have different estimates of the normal yield and, as a result the critical yield, one can finally construct an \textit{aggregated} demand for speculative balances (Fig. 2.3b). The resulting aggregated money demand curve has a continuous shape with a negative slope, since with a decreasing market interest rate more and more investors will expect an increase of this rate and, therefore, prefer to hold money instead of bonds.

\[\text{(a) Individual demand for speculative money balances} \quad \text{(b) Aggregated demand for speculative money balances}\]

\[r_{\text{critical}} \quad \text{Demand for speculative balances, } L_s\]

\[r \quad \text{Demand for speculative balances, } L_s\]

\textit{Fig. 2.3} Keynesian demand for speculative balances
In our example, assume that the normal yield is 10% p.a. According to Eq. (2.20b), the critical market yield is:

\[ r_{\text{critical}} = \frac{0.1}{1 + 0.1} \approx 9\%. \] (2.20c)

So if the market yield is less than 9%, the investor wouldn’t want to hold any bonds as expected capital losses are higher than interest income.

It is important to note that in the Keynesian liquidity theory, the speculative demand for money always equals the supply of bonds and vice versa.\(^4\) As it is important for the understanding of the portfolio approach underlying the Keynesian theory, this aspect shall be explained in some more detail.

In a first step, let us assume the economy’s number of bonds outstanding, \(b_0\), can be traded at any time. Moreover, the demand for bonds, \(B_D\), which is assumed to be independent of the bond price, amounts to US$100. Given a bond supply of \(b_0 = 1\) (for instance, 1 million bonds outstanding), the equilibrium price is US$100 (Fig. 2.4a). Alternatively, would the number of offered bonds increase to, say, \(b_1 = 2\), the equilibrium price of bonds, \(P_B\), would fall to US$50.

In a second step, we show the relation between bond prices (vertical axis) and the market capitalisation of bonds (horizontal axis), that is \(b\) times \(P_B\) (Fig. 2.4b). The economy’s total bond demand, \(B_D = US$100\), is assumed to be given, represented by the vertical line. Bond supply, \(BS_0\), is represented by a line with a 45° slope (\(tg \alpha = 1\) if \(b_1 = 1\)); assuming \(b_1 = 2\), the higher bond supply would be represented by \(BS_1\) (with \(tg \alpha = 0.5\)), since the market volume would double in this case (when compared to \(BS_0\)).

---

\(^4\)See Borchert (2001), pp. 120.
In a third step, we replace the bond price on the vertical axis of Fig. 2.4b by the bond yield $r$, where:

$$r = \frac{NV \cdot n}{P_B},$$

with $NV =$ nominal value of the bond, $n =$ coupon and $P_B =$ bond market price.

In view of the pricing formula above, the bond demand curve stays the same as before, and the bond supply curve converges asymptotically towards the vertical and horizontal axis (Fig. 2.5). The higher bond price $P_B = 100$ in Fig. 2.4(b) corresponds with the lower bond yield $r_0$ in Fig. 2.5, and the lower bond price $P_B = 50$ with the higher bond yield $r_1$. In Fig. 2.5, bond demand equals money supply (and vice versa), that is $B_D = M_S$, which is representing the demand for speculative balances.

That said, a given liquidity preference function is actually defined on the basis of a given stock of bonds and a given stock of money, whereas bonds and money represent the economy’s total stock of wealth, $W$:

$$L_s = L_s(r, W), \text{ mit } \partial L_s/\partial W > 0. \quad (2.21)$$

An increase in the supply of bonds would move the liquidity preference curve upwards: under a given stock of money, an increase in the supply of bonds would be associated with a higher interest rate. So any open market operation by the central bank, for instance, which brings about a change in money supply, would also affect the location of the liquidity preference function.

---

If the above bond pricing formula is valid and the price of the bonds $P_B$ converges towards zero, the interest rate $r$ tends to infinity and vice versa. Hence, the curve depicting the bond supply reveals the shape described above.
2.2 Keynesian Money Demand Theory

2.2.1 Explaining the Trend of Income Velocity of Money

The Keynesian demand for money theory can be integrated in the quantity theory framework for analysing the behaviour of the income velocity of money. The long-run demand for real money balances is:

\[ m_t - p_t = \beta_0 + \beta_1 y_t - \beta_2 r_t + \varepsilon_t, \tag{2.22} \]

where \( y_t \) is the log of real income in period \( t \), \( \beta_0 \) is a constant term, \( \beta_1 \) and \( \beta_2 \) represent the income and interest rate elasticities of money demand, respectively, and \( r_t \) is the interest rate. If \( \varepsilon_t \) is a stationary stochastic process with zero mean, Eq. (2.22) describes a cointegrated long-run relationship (King & Watson, 1997).

Combining equation (2.22) with the log version of the transaction equation, that is \( m - p = y - \nu Y \), yields the following expression for the income velocity of money:

\[ \nu_{y,t} = -\beta_0 + (1 - \beta_1) \cdot y_t + \beta_2 \cdot r_t - \varepsilon_t. \tag{2.23a} \]

As far as \( r_t \) is concerned, it would be appropriate for a broadly defined monetary aggregate to use the difference between the opportunity cost of money holdings and the (weighted) yield paid on deposits included in the definition of money. If the yield spread is stationary, however, the interest component in (2.23a) could not be held responsible for the velocity of money’s trend behaviour. Under this assumption, the functional specification of the demand for money would actually be tied to current income as the interest rate \( r \) drops out:

\[ \nu_{y,t} = -\beta_0 + (1 - \beta_1) \cdot y_t - \varepsilon_t. \tag{2.23b} \]

If the log velocity in time \( t \) fluctuates randomly around a constant term, it becomes \( \nu_{y,t} = \nu_0 + \varepsilon_t \). With \( \varepsilon_t \) being a stationary zero mean process, the equilibrium level of log velocity would be \( \nu^* = \nu_0 \). Hence, calculating trend velocity finally yields:

\[ \nu^* = \nu_0 + (1 - \beta_1) \cdot y^*, \tag{2.24} \]

with \( y^* \) log of potential output. If, for instance, \( \beta_1 \) exceeds unity, the income velocity of money would predict a declining trend over time as long as potential output is growing (assuming throughout that the opportunity costs of money holdings is integrated of order zero (I(0))).

2.2.2 Some Empirically Testable Money Demand Hypotheses

In view of the findings above, there are some testable relationships which have become important in empirical analyses of money demand functions (Herwartz &
Reimers, 2006). To start with, one may be concerned with cointegration between money, prices and real output, based on the static representation of real money, which corresponds to:

\[ m_t - p_t = \beta_0 + \beta_1 y_t + z_t, \]  \hspace{1cm} (2.25)

excluding interest rates for reasons stated above, and with \( z_t \) being a stationary stochastic process with zero mean. To test for cointegration, the potential cointegrating relationship may be written to allow the interpretation of a price equation:

\[ p_t = -\beta_0 - \beta_1 y_t + \beta_2 m_t - z_t. \]  \hspace{1cm} (2.26)

The first hypothesis to consider is then:

\( H_1: p_t, m_t \) and \( y_t \) are cointegrated with cointegration rank \( r = 1 \).

Second, deviations from the long-run relation should not affect the change in output:

\( H_2: Output \) is weakly exogenous.

Third, the coefficients of prices and money should be equal in absolute terms:

\( H_3: \beta_2 = 1 \) which suggests price homogeneity.

Fourth, if the cointegration relation is normalised such that the coefficient of the price level is unity, the coefficient of real output is unity:

\( H_4: \beta_1 = 1, \)

implying that the velocity of money is constant in the long-run. If \( \beta_1 > 1 \), trend income velocity follows a downward trend over time.

The Keynesian money theory has had a profound impact. Most importantly, it has ended the dichotomy the (neo-) classical theory assumed between the economy’s real and monetary sector. Under the assumption that investment spending and the demand for money depend on the interest rate, the central bank, by changing the stock of money and thereby affecting the interest rate, was supposed to exert an impact on aggregate demand. Over time, elements of the Keynesian money demand system have been developed further, particularly in portfolio-oriented money demand theories.

### Overview of Functional Forms of Money Demand

Table 2.1 summarizes the salient features of various functional forms which are now widely used for estimating money demand functions.

The elasticity \( E \) of a variable \( M \) (such as, for instance, money) with respect to another variable \( X \) (say, the interest rate) is defined as:
### Table 2.1 Functional forms of money demand

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
<th>Slope (\frac{dY}{dX})</th>
<th>Elasticity (\frac{dY}{dX} \cdot \frac{X}{Y})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>(Y = \beta_1 + \beta_2 X)</td>
<td>(\beta_2)</td>
<td>(\beta_2 \left(\frac{X}{Y}\right))^*</td>
</tr>
<tr>
<td>Log-linear or</td>
<td>(\ln Y = \beta_1 + \beta_2 \ln X)</td>
<td>(\beta_2 \left(\frac{\ln X}{\ln Y}\right))</td>
<td>(\beta_2)</td>
</tr>
<tr>
<td>Log-log</td>
<td>(\ln Y = \beta_1 + \beta_2 X)</td>
<td>(\beta_2(Y))</td>
<td>(\beta_2(X)^*)</td>
</tr>
<tr>
<td>Lin-log</td>
<td>(Y = \beta_1 + \beta_2 \ln X)</td>
<td>(\beta_2 \left(\frac{1}{X}\right))</td>
<td>(\beta_2 \left(\frac{1}{Y}\right)^*)</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>(Y = \beta_1 + \beta_2 \left(\frac{1}{X}\right))</td>
<td>(-\beta_2 \left(\frac{1}{X^2}\right))</td>
<td>(-\beta_2 \left(\frac{1}{XY}\right)^*)</td>
</tr>
</tbody>
</table>

Note: An asterisk indicates that the elasticity coefficient is variable, that is depending on the value taken by \(X\) or \(Y\) or both.

\[
E = \frac{\% \text{change in } Y}{\% \text{change in } X}
\]

\[
E = \frac{(\Delta Y/Y) \cdot 100}{(\Delta X/X) \cdot 100}
\]

\[
E = \frac{\Delta Y \cdot X}{\Delta X \cdot Y}
\]

\[
E = \text{slope} \cdot (X/Y),
\]

where \(\Delta\) denotes a (small) change. If \(\Delta\) is sufficiently small, it can be replaced by the calculus derivate notation \(dY/dX\).

Take, for instance, the following money demand function:

\[
\ln M = \beta_1 + \beta_2 \ln Y + \beta_3 \ln r + e, \quad (2.27)
\]

where \(M\) = (real) stock of money, \(Y\) = real output and \(r\) = interest rate and \(e = \text{i.i.d. error term}\). This equation can also be estimated using the level instead of the logarithm of the interest rate:

\[
\ln M = \beta_1 + \beta_2 \ln Y + \beta_3 r + e. \quad (2.28)
\]

In Eq. (2.28), the coefficient of \(r\) represents the semi-elasticity. It shows by how much percent the real demand for money changes in response to a change in the interest rate of 1 percentage point (that is, for instance, the rate rising from 5% to 6%).
In contrast, Eq. (2.27) shows how the demand for money changes in response to a given percentage point change in the interest rate (say, a one percent change from 5.0% to 5.05%). That said, the semi-elasticity is defined as:

$$\beta_2 = \frac{d \ln M}{dr} = \frac{dM/M}{dr}.$$  (2.29)


### 2.3 Portfolio Oriented Money Demand Theory

#### 2.3.1 Monetarist Money Demand

Milton Friedman (1956) extended Keynes’ speculative money demand within the framework of asset price theory, treating money as a good as any other durable consumption good. To Friedman, the demand for money is a function of a great number of factors. His analysis refers to nominal magnitudes in the first place. Taking this as a starting point, he derives the real money demand from nominal money demand under certain assumptions.

Perhaps most important, Friedman maintains that it is permanent income – and not current income as in the Keynesian theory – that determines the demand for money. If $W^n$ represents market agents’ total wealth, which comprises all sources of income – among them human capital and all consumable services –, and $r$ is the representative market yield, nominal income per period is: \(^6\)

$$Y^n = W^n \cdot r.$$  

With $W^n$ and $r$ assumed to be relatively stable, market agents’ permanent income can be also expected to be rather stable when compared with current income.

More specifically, Friedman’s monetarist demand function for money can be summarized as follows (Felderer & Homburg, 2005, p. 244):

$$L^n = f \left( P^{(+)}, r_B^{(-)}, r_E^{(-)}, \frac{\dot{P}}{P^{(-)}}, \frac{Y^n}{r^{(+)}} \right).$$  (2.30)

\(^6\)Such a specification can be interpreted as an income or wealth hypothesis. See Meltzer (1963), p. 220.
According to Eq. (2.30), the nominal demand for money:

- is positively related to the price level, \( P \), as market agents are expected to hold a certain stock of real rather than nominal money balances. If the price level rises (falls), the demand for money increases (decreases);
- rises (falls) if the opportunity costs of money holdings – that are bonds and stock returns, \( r_B \) and \( r_E \), respectively – decline (increase);
- is influenced by inflation, \( \dot{P}/P \) (or: \( dP/dt \)). A positive (negative) inflation rate reduces (increases) the real value of money balances, thereby increasing (reducing) the opportunity costs of money holdings. Whereas a *one-off* increase in the price level would increase the demand for money, inflation – that is an ongoing increase in the economy’s price level – will lead to a decline in money demand.
- is a function of total wealth, which is represented by permanent income divided by the discount rate \( r \), defined as the average return on the five asset classes in the monetarist theory world, namely money, bonds, equity, physical capital and human capital, \( \upsilon \) (which is, for the sake of clarification, explicitly shown in Eq. (2.30)).

Furthermore, Friedman assumes that \( r \), which cannot be measured directly, would correspond more or less with \( r_B \) and \( r_S \), so that \( r \) (and \( \upsilon \)) can be dropped from the equation:

\[
L^n = f \left( P, r_B, r_E, \frac{\dot{P}}{P}, Y^n \right). \tag{2.31}
\]

If market agents do not suffer from *money illusion*, a change in \( P \) and a change in \( Y^n \) by the multiple \( \lambda \) will change the demand for money by the same amount (*homogeneity hypothesis of money demand*):

\[
\lambda \cdot L^n = f \left( \lambda \cdot P, r_B, r_E, \frac{\dot{P}}{P}, \lambda \cdot Y^n \right). \tag{2.32}
\]

If (2.32) holds true for any arbitrary realisation of the parameter \( \lambda \), we can, for instance, define \( \lambda = 1/Y^n \). Substituting this expression in (2.32) yields:

\[
\frac{1}{Y^n} \cdot L^n = f \left( \frac{P}{Y^n}, r_B, r_E, \frac{\dot{P}}{P}, 1 \right) \text{ and, by rearranging,} \tag{2.33}
\]

\[
L^n = f \left( \frac{P}{Y^n}, r_B, r_E, \frac{\dot{P}}{P}, 1 \right) \cdot Y^n. \tag{2.34}
\]

According to Eq. (2.34), Friedman’s demand for money is a stable function (i.e., \( f(.) \) times \( Y^n \)) of several variables – rather than a numerically constant number as is assumed in the (neo-)classical theory. The velocity of money \( v (= Y^n/L^n) \) is the reciprocal of \( f(.) \), which, according to (2.34), is equal to \( L^n/Y^n \). That said, we can write (dropping the “1”):
\[ v \left( \frac{Y^n}{P}, r_B, r_E, \frac{\dot{P}}{P} \right) = \frac{1}{f \left( \frac{P}{Y^n}, r_B, r_E, \frac{\dot{P}}{P}, 1 \right)}. \] (2.35)

Imposing market equilibrium, \( M = L^n \), we arrive at the new formulation of the equation of exchange:

\[ M \cdot v \left( \frac{Y^n}{P}, r_B, r_E, \frac{\dot{P}}{P} \right) = Y^n, \] (2.36)

which actually restates the traditional specification of the quantity theory.

In this context it seems interesting to highlight the effect a one-off money supply increase might exert on the real economy from the point of view of the Monetarists. With the economy running at full capacity, a one-off increase in money supply of, say, 5%, can be expected to increase (future) inflation. This, in turn, would reduce peoples’ demand for money, thereby increasing the velocity of money \( v = Y^n/M \) by, say, 2%.

As a result, Friedman’s money demand function suggests that the percentage increase in nominal income should be higher than a given percentage increase in money supply, namely 7% in our example (that is 5% money growth plus a 2% increase in velocity). However, as soon as the additional money supply has fully translated into an increase of the price level, inflation returns back to its previous level, increasing the demand for money, thereby lowering the velocity of money by 2%. Finally, after all adjustments have run their course, nominal output is expected to have risen by 5%, corresponding to the increase in the stock of money. To sum up, Monetarists would expect that a one-off increase in the money supply would exert cyclical swings of the economy’s nominal income.

### 2.3.2 Post-Keynesian Money Demand Theory

Two characteristics of money provide the starting point for Post-Keynesian theories of money demand. First, the analysis of the transaction function of money has brought forward inventory models. Second, the study of the store-of-value function of money has inspired asset or portfolio model approaches, in which money is held as part of market agents’ overall asset portfolios. Our discussion begins with the inventory model approach to transaction balances.

#### 2.3.2.1 Inventory Model Approach to Transaction Balances

Baumol (1952) and Tobin (1956) developed a deterministic theory of money demand in which money was essentially viewed as an inventory held for transaction purposes. Although relatively liquid financial assets other than money (such as, for instance, time and savings deposits) offer a positive return, the transaction costs of going between money and these assets justifies holding money. Inventory models actually assume two media for storing value: money and an interest
bearing alternative asset. What is more, inventory models state that there is a fixed cost of making transfers between money and the alternative asset. Finally, it is assumed that all transactions involve money as the means of exchange, and that all relevant information is known with certainty.

Let us assume that a household receives a nominal income of $PY$ in each period (say, a month), spending if at a constant rate over the period. Money which is not needed for purchasing will be put in an interest-bearing account. As goods are purchased with cash, the household needs to go to his bank at least once (for converting sight deposits into cash). Each visit to the bank has a nominal cost $c$ attached to it, and $c$ should be thought of as the opportunity cost (such as time spent queuing, shoe leather costs, fees for using online banking services, etc). If $n$ is the number of visits to the bank during the month, the monthly cost will be: $n \cdot c$. To reduce costs, the household might wish to hold more cash. However, holding more cash comes at a cost in the form of foregone interest income. If $i$ is the monthly interest rate on the savings account (the alternative to cash), holding an average nominal balance $M$ over the month gives an opportunity cost of: $i \cdot M$. Figure 2.6 illustrates the average transaction balance holdings as a function of $n$. $PY$ denotes the incoming payment per month. If, for instance, the household spends his money continuously over the period, his money account is zero at the end of the month.

For evenly-spaced visits to the bank, the average money balance is $PY/n$. These $n$ trips occur every $1/n$th of a month. The area of each triangle represents the average

![Fig. 2.6 Number of bank trips and average money holdings](image-url)
money amount held between two trips. That said, the average money holdings as a function of the number of bank trips, \( n \), are:

\[
\text{average money holdings between } n \text{ bank trips, } M = \frac{PY}{2n}. \quad (2.37)
\]

The opportunity cost is the foregone interest on the average money holdings:

\[
\text{opportunity costs} = M \cdot i. \quad (2.38)
\]

Transaction costs are a function of the transaction costs per bank trip, \( c \), and \( n \):

\[
\text{transaction costs} = c \cdot n. \quad (2.39)
\]

Total costs, \( TC \), taking into account the opportunity costs of money holdings and transaction costs, are therefore:

\[
TC = M \cdot i + cn. \quad (2.40a)
\]

As \( n = \frac{PY}{2M} \) (remember that the average money holdings between \( n \) bank trips amount to \( M = \frac{PY}{2n} \)), one can write:

\[
TC = M \cdot i + c \cdot \frac{PY}{2M}. \quad (2.40b)
\]

The optimization of money holdings is equivalent to:

\[
TC = M \cdot i + c \cdot \frac{PY}{2M} \rightarrow \min. \quad (2.40c)
\]

The minimum costs of money holdings are obtained by taking the first derivative of equation (2.40c) with respect to \( M \), setting the result to zero:

\[
\frac{\partial TC}{\partial M} = i - \frac{cPY}{2M^2} = 0. \quad (2.41)
\]

Since the sufficient condition of a cost minimum is fulfilled (note that the realization of the second derivation is positive), the optimal holding of transaction balances is the following \textit{square root formula}:

\[
M^* = \sqrt{\frac{cPY}{2i}}. \quad (2.42)
\]

Expression (2.42) states that the optimal average money holding is:

- a positive function of real economic activity \( Y \),
- a positive function of the price level \( P \),
– a positive function of transactions costs $c$, and
– a negative function of the nominal interest rate $i$.

If we define the real cost of transactions, $c^{\text{real}}$, as $c^{\text{real}} = c/P$, the square root formula can be expressed in terms of real money demand:

$$\left( \frac{M}{P} \right)^* = \sqrt{\frac{c^{\text{real}} Y}{2i}}. \quad (2.43)$$

The real demand for transaction balances in natural logarithms, $m - p$, can be written as:

$$m - p = 0.5 \ln \left( \frac{c^{\text{real}}}{2} \right) + 0.5 \ln Y - 0.5 \ln i. \quad (2.44)$$

It is determined by the constant $c^{\text{real}}/2$, real income and the interest rate. Accordingly, the income elasticity of money is:

$$\varepsilon_{m-p, Y} = \frac{d(M/P)}{M/P} ; \frac{dY}{Y} = \frac{d \ln(m - p)}{d \ln Y} = 0.5. \quad (2.45)$$

The interest rate elasticity of money demand can be written as:

$$\varepsilon_{i,Y} = \frac{d(M/P)}{M/P} ; \frac{di}{i} = \frac{d \ln(m - p)}{d \ln i} = -0.5. \quad (2.46)$$

To conclude, in contrast to the Keynesian demand for transaction balances, the transaction demand of Baumol and Tobin is also a function of interest rates. What is more, the income elasticity suggests economies of scale of transaction money holdings. A rise in real income by, say, 1 percentage point would, according to Tobin and Baumol, be followed by a 0.5 percentage point increase in real balance holdings – rather than a proportional increase as suggested by the transaction equation.7

2.3.2.2 Tobin’s Demand for Speculative Balances

Tobin’s mean-variance analysis of the demand for speculative money holdings is in fact an application of the theory of portfolio choice. Tobin (1958) assumes that the utility which people derive from their asset holdings is positively related to the expected return of the portfolio and negatively related to the risk of the asset portfolio, with the latter being measured by the variance (or standard deviation) of asset returns.

An individual’s preference for return and risk can be illustrated by means of indifference curves as shown in Fig. 2.7. In the expected return-risk-space, the indif-

---

7 The approach to money demand derived in this section may, for instance, be also applied to professional cash managers in large firms and, in particular, to investment and commercial banks’ management of base money holdings.
ference curves slope upward because an individual is willing to accept more risk if it is accompanied by a higher expected return. When we turn to higher indifference curves, utility is higher because for the same level of risk, the expected return is higher. Against this background, we now take a closer look at the mean-variance model for explaining the demand for speculative balances.

To start with, we assume that total wealth, $W$, consists of the stock of money, $M$, and the stock of bonds, $B$, that is:

$$W = M + B,$$  

(2.47)

where $B$ represents the market value of bonds, that is the bond price, $P_X$, times the number of outstanding bonds, $b$. The expected return on a bond portfolio, $R^e$, is determined by the interest payment, $r$, and the expected capital gain, $g$:

$$R^e = r \cdot B + g \cdot B = (r + g) \cdot B.$$  

(2.48)

In contrast to Keynes, expected capital gains are no longer assumed to be known with certainty but characterized by a measure of the distribution of possible returns around the mean value:

$$g \sim N(\bar{g}, \sigma_g^2),$$

where the expected capital gain is $E(g) = \bar{g}$ and the variance of the capital gain is $\text{var}(g) = \sigma_g^2$.

If all bonds are assumed to have the same risk, $\sigma_g$, the total portfolio risk is:

$$\sigma_P = \sigma_g \cdot B.$$  

(2.49)
Solving (2.49) for $B$ and inserting in (2.48) yields the *transformation curve*, showing the opportunity locus of all possible combinations of $R^e$ and $\sigma_p$ for alternative asset allocations:

$$R^e = \frac{r + \bar{g}}{\sigma_g} \cdot \sigma_p. \quad (2.50)$$

The utility function of a *risk averter* can be written as follows:

$$U = U(R^e, \sigma_p), \quad \text{where} \quad \frac{\partial U}{\partial R^e} > 0, \quad \frac{\partial U}{\partial \sigma_p} < 0, \quad \frac{\partial^2 U}{\partial^2 \sigma_p} > 0. \quad (2.51)$$

Note that the last inequality implies that the utility loss induced by risk becomes increasingly larger – for a given level of return.

The optimal portfolio can be calculated by applying the *Lagrange technique*:

$$L = U(R^e, \sigma_p) + \lambda \left( R^e - \frac{r + \bar{g}}{\sigma_g} \sigma_p \right) \rightarrow \text{max}. \quad (2.52)$$

Differentiating (2.52) for $R^e$ and $\sigma_p$, setting the first derivations to zero and solving for $\lambda$ yields:

$$\lambda = -\frac{\partial U}{\partial R^e} \quad \text{and} \quad \lambda = -\frac{\partial U}{\partial \sigma_p} \cdot \frac{\sigma_g}{r + \bar{g}}. \quad (2.53a, b)$$

Equating (2.53a) and (2.53b) and rearranging yields the optimum condition:

$$-\frac{\partial U}{\partial R^e} = \frac{\partial U}{\partial \sigma_p} \cdot \frac{\sigma_g}{r + \bar{g}} \quad \text{or} \quad \frac{\partial U}{\partial R^e} \cdot (r + \bar{g}) = \frac{\partial U}{\partial \sigma_p} \cdot \sigma_g. \quad (2.54)$$

Equation (2.54) shows that, in the optimum, the slope of the indifference curve is equal to the slope of the transformation curve: the increase in utility from holding an additional bond unit $[(r + \bar{g}) \cdot (\partial U / \partial R^e)]$ equals the decline in utility from taking additional risk $[\sigma_g \cdot \partial U / \partial \sigma_p]$.

To calculate the actual portfolio structure, we make use of the total wealth restriction (2.47): $\bar{W} = M + B$. Solving for $M$ and substitution $B$ by the term $\sigma_p / \sigma_g = B$, we can write:

$$M = \bar{W} - \frac{1}{\sigma_g} \sigma_p, \quad (2.55)$$

which shows the *demand for money* as a function of total wealth, the individual risk of bonds and the total bond portfolio risk.

That said, in Tobin’s theory of the demand for speculative balances individuals are capable of holding money as well as bonds (mixed portfolios). This is different
from Keynes and his liquidity preference theory, in which an individual holds either money or bonds; in the liquidity preference theory, the possibility of holding mixed portfolios is limited to the macroeconomic level.

The results are illustrated in Fig. 2.8. Quadrant 1 shows the demand for speculative balances as a negative function of expected return insert eq. (2.50) in (2.55). Quadrant 2 depicts the transformation curve (Eq. 2.50), represented by $T_0$. Quadrant 3 plots the relation between risk and bond holdings.

At point $C(U_0)$ the slope of the indifference curve equals the slope of the transformation curve, representing the investor’s optimum portfolio choice. Here, the bond holding is $B(r^*, \sigma_g)$. The optimal money holding is $L_C$, corresponding to point $C'$. In the following, the implications of three scenarios for the demand for money and bond holdings shall be analyzed, namely a change in the investor’s estimate of (i) return, (ii) risk and (iii) a change in taxes.

### 2.3.2.3 Re (i): Change in Expected Return

Let us assume that – starting from point $C(U_0)$ – the bond yield rises from $r^*$ to $2r^*$. As a result, the transformation curve steepens (from $0T_0$ to $20T_1$). The new optimum is in point $D(U_1)$. Bond holdings increase to $B(2r^*, \sigma_g)$ – with the risk-bond relation remaining unaffected –, giving the new risk-bond relation at point $D''$ in quadrant 3. Accordingly, with a growing share of bonds in the portfolio, the demand for money balances declines – along the liquidity preference curve in quadrant 1 $L(\sigma_g)$ – from $L_C$ to $L_{D'}$. That said, a given liquidity preference curve reflects different levels of return expectations under a given risk perception.
2.3.2.4 Re (ii): Change in Risk

Starting with the optimum portfolio in $C(U_0)$, the assumption is that investors cut their risk estimate in half, from $\sigma_g$ to $\sigma_g/2$. As a result, the transformation curve in quadrant 2 will double in slope, from $OT_0$ to $OT_1$. Consequently, the new optimum point $D(U_1)$ corresponds to bond holdings of $B(r^*, \sigma_g/2)$, now expressed by point $D''$ in quadrant 3. With a higher share of bonds in the optimal portfolio, an unchanged expected return of $r^*$ implies a corresponding decline in liquidity preference from $L_C$ to $L_D''$. The liquidity preference curve becomes steeper, leading to this new equilibrium $D''$ in quadrant 1. That said, the change in risk perception entails a change in the interest rate elasticity of the liquidity preference curve (note that an unchanged demand for transaction balances is assumed). Generally speaking, the Tobin model would suggest that for a given expected return the interest rate elasticity of speculative money demand declines (rises) when the risk perception declines (increases).

Heightened Investor Uncertainty and Money Holdings

The period between 2000 and 2003 provides an illustration of how changes in investor risk and return perception can (temporarily) influence the demand for money – as suggested by Tobin’s theory of the demand for speculative balances. The early years of the 21st century were characterised by heightened financial market uncertainty. The key events include the bust marking the end of the IT-driven “New Economy” boom, the terrorist attacks in the US on 11 September 2001, a spate of accounting scandals on both sides of the Atlantic in the aftermath of the equity market correction, and the start of the wars in Afghanistan in late 2001 and in Iraq in early 2003. All these events contributed in one way or another to a significant and protracted fall and heightened volatility in global stock prices from mid-2000 onwards.

As a consequence of strongly increased investor risk perception, extraordinary portfolio shifts affecting monetary aggregates took place. These shifts strongly influenced money growth in many countries in a way that could not be easily explained by the conventional determinants of money demand, such as prices, income and interest rates.

At the global level, precautionary and speculative motives significantly influenced the overall demand for money during that period. The fact that common global shocks might have influenced the demand for money in several regions was illustrated by a rather close co-movement of the growth rates of broad monetary aggregates during the period of heightened financial market turmoil.

Figures 2.9(a to d) show annual money supply growth rates and measures of stock market volatility in the US, the euro area, Switzerland and the UK for
2.3.2.5 Re (iii): Increase in Taxes

Suppose the initial position of the investor is $D(U_1)$ on the transformation line $0T_1$, with bond holdings of $B(2r^*, \sigma_g)$ and money demand at $L_D$. Now, investors cut in half expected returns (from $2r^*$ to $r^*$) and perceived risk (from $\sigma_g$ to $\sigma_g/2$). The

---

8 In the case of the US, Switzerland and UK, stock market volatility is expressed by the market traded volatility index of the S&P’s 500; in the case of the euro area it is the V-DAX.
transformation curve will remain unchanged at \( 0T_1 \), and investors will still wish to obtain the combination of risk and reward depicted by \( D(U_1) \). This combination, however, requires investors to double their holdings of bonds, that is moving to \( B(r^*, \sigma_g/2) \); the tax change moves the risk-consol relation from \( 0Q_0 \) to \( 0Q_1 \). Accordingly, the demand for cash declines from \( L_{D'} \) to \( L_{D''} \).

**Digression: Income Velocities of US Monetary Aggregates**

Figures 2.10 (a–d) show the income velocities of US monetary aggregates – that is M1, M2, M2-ST (the latter being M2 minus time deposits) – for the period 1960-Q1 to 2008-Q1 and the income velocity of M3 up to 2006-Q1. Two findings stand out. First, the income velocities for all aggregates are *not numerical constants*, as suggested by the (neo)classical quantity theory. Second, the income velocities tend to exhibit rather *pronounced swings* over time.

---

**Fig. 2.10** Income velocities of US monetary aggregates  
Source: Bloomberg, Federal Reserve Bank of St. Louis; own calculations. – The velocities were calculated by dividing nominal GDP by the stock of money. – M2-ST is M2 minus time deposits. – Period: 1960-Q1 to 2008-Q1 (M3: until 2006-Q1)
Take, for instance, the income velocity of the stock of M1. Its trend pointed upwards from the late 1960s until around 1981-Q1. From then on until the end of 1994, the trend of income velocity of M1 declined, and started rising thereafter. In the period under review, the income velocity of M2 exhibited an \textit{upward trend}, accompanied with a rather strong rise in the early 1990s. By contrast, the income velocity M3 suggests a long-term \textit{downward trend}.

Time-varying income velocities of money do not necessarily indicate \textit{unstable} income velocities of the money. In fact, the crucial question is whether a monetary aggregate’s income velocity – or, its reciprocal, the demand for money – is a reliable function of various observable variables such as, for instance, output, prices, risk perception, and interest rates. In that sense, changes in the income velocity would reflect changes in its determining factors.

From the mid-1980s to the early 1990s, the income velocity of M2 was perceived to be reasonably stable and the most reliable of the alternative money measures. M2 and nominal GDP had grown at approximately the same rate over the previous 30 years, suggesting a simple and robust relationship that provided monetary policy makers with an uncomplicated framework for setting monetary targets. This relationship was summarized by the trendless long-run average of M2 velocity up to the end of the 1980s.

Movements of actual income velocity of M2 away from this level were \textit{positively correlated with the opportunity costs} of M2 holdings, measured as the spread between the 3-months Treasury Bills rate and the weighted average returns on deposits included in the stock of M2 (Fig. 2.11). When short-term interest rose, opportunity costs of M2 holdings increased because the rates on components of M2 did not climb as fast as market interest rates.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2_11.png}
\caption{Income velocity and opportunity costs of M2}
\end{figure}

\textit{Fig. 2.11} Income velocity and opportunity costs of M2
Source: Federal Reserve Bank of St. Louis, Bloomberg; own calculations. – The income velocity of the monetary aggregate was calculated by dividing nominal GDP by the stock of M2. – Period: 1970-Q1 to 2005-Q4. – Opportunity costs of M2 holdings is calculated by the 3-months Treasury Bills rate minus the weighted average return of deposits included in M2.
In the first half of the 1990s, however, income velocity of M2 rose substantially above its long-term trend level, despite a considerable decline in M2 opportunity costs (Hallman & Anderson, 1993). Conventional M2 demand function, which modelled the relationship between M2 and variables such as output and interest rates began to go off track. As the link between M2 and GDP deteriorated, the forecasting ability of M2 money demand equations suffered. Friedman and Kuttner (1992) argued that by the early 1990s the relationship between M2 and GDP had weakened, and in July 1993 Federal Reserve Chairman Alan Greenspan reported that: “(...) at least for the time being, M2 has been downgraded as a reliable indicator of financial conditions in the economy, and no single variable has yet been identified to take its place.”

The difficulties in forecasting M2 spurred a fair amount of research examining whether the deterioration in the M2 equation’s forecasting ability was temporary, or whether more fundamental factors – such as flaws in the construction of the opportunity cost, the M2 aggregate, or both – were at work. Carlson and Parrott (1991), and Duca (1992) argued that the existence of troubled US thrift institutions and the length of time it took the Resolution Trust Corporation to resolve the thrifts’ difficulties helped explain the weakness in M2.

Duca found that the change in the volume of cumulated deposits at resolved thrifts accounted for a large part of the decline in M2 growth. Orphanides, Reid and Small (1994) and Duca (1995) examined whether some of the weakness in M2 growth reflected a substitution by households away from M2-type deposits and into bond and equity mutual funds. They found that this substitution effect appeared to account for only a small part of the M2 weakness. Collins and Edwards (1994) tried to restore the hitherto stable demand function by redefining M2 (that is including additional instruments such as bond mutual funds in M2).

Carlson and Keen (1995) suggest that preliminary evidence indicates that the relationship between M2, inflation, and output may have stabilized. The prediction errors of the M2 demand model have essentially remained unchanged since 1992. This is consistent with a permanent one-time shift in the level of M2 relative to income. Such an outcome would be the case if the forces underlying the deceleration of M2 – such as, for instance, restructuring of credit markets and financial innovation – had worked themselves out.

Koenig (1996a, b) notes that attributing the slowdown in M2 growth to the thrift resolution process had largely been abandoned in dealing with the demand for M2 function. The focus instead has shifted to an examination of the competitiveness problems of financial intermediaries in the face of tighter regulations and stricter capital standards. Koenig proposes an alternative approach by altering the opportunity cost measure to include a long-term Treasury bond rate in the M2 demand function.

---

9The work of Estrella and Mishkin (1997) provided further support for this finding.
Lown, Peristiani and Robinson (1999) argue that the financial conditions of depository institutions were a major factor behind the unusual pattern of M2 growth in the early 1990s. By constructing alternative measures of M2 based on banks’ and thrifts’ capital positions, the authors show that the anomalous behavior of M2 in the early 1990s disappears. After accounting for the effect of capital constrained institutions on M2 growth the unusual behavior of M2 velocity during the early 1990s can be explained, and a more stable relationship between M2 and the ultimate goals of policy can be established. Lown and Peristiani conclude that M2 may contain useful information about economic growth during periods when there are no major disturbances to depository institutions.

Following Carlson and Keen (1995), the income velocity of M2 shall be estimated on the basis of a level shift variable which takes account of the rise in income velocity starting in the early 1990s. The shift parameter takes the value of 0 from 1970-Q1 to 1989-Q4, then rises linearly to 1 until 1994-Q4 and remains constant thereafter. The estimation result is shown in Fig. 2.12(a), which depicts the actual and estimated income velocity of M2 and the deviation from mean (residual).

Following the adjustment suggested by Carlson and Keen (1995), the historical relation between deviations of income velocity from its long-run trend and M2 opportunity costs can actually be restored (Fig. 2.12b). One implication of this result might be that M2 velocity has stabilized at a higher level, while the short-run relationship between swings in the income velocity of M2 and the opportunity cost of M2 holdings has remained basically unchanged.

Fig. 2.12  Actual and estimated income velocity of M2
Source: Bloomberg; own calculations. – With Carlson and Keen (1995), the trend income velocity is estimated to 0 from 1970-Q1 to 1989-Q4, thereafter rising linearly to 1 until 1992-Q4 and remaining constant at that level. – The regression is: \( VM_2 = a + TL_t + \varepsilon_t \), where \( VM_2 \) = income velocity of money, \( TL \) takes account of the level shift in the income velocity of M2, \( \varepsilon_t \) is i.i.d. “white noise” error term. – Period 1968-Q1 to 2005-Q4.
2.4 Money-in-the-Utility Function and Cash-In-Advance Models of Money Demand

An alternative theoretical explanation of why people wish to hold money can be found in money-in-the-utility-function (MUF) and cash-in-advance (CIA) theories of money demand. Both approaches do justice to the notion that people derive utility from holding money.

2.4.1 Money-in-the-Utility Function of Money Demand

The utility function of a representative market agent is:

\[ U_t = \sum_{s=t}^{\infty} \beta^{s-t} u \left( C_s, \frac{M_s}{P_s} \right), \quad (2.56) \]

where \( \beta \) is the discount factor, \( C_t \) real consumption, \( M_t \) nominal stock of money (acquired at the beginning of period \( t \) and held to the end of \( t \)), and \( P \) is the price level. We assume further that the marginal utility of consumption and real money holdings is positive, and that the utility function is strictly concave.

What is more, it is assumed that in a one-good-only open economy the purchasing power parity (PPP) holds:

\[ P_t = \varepsilon_t P^*_t, \quad (2.57) \]

where \( \varepsilon \) represents the price of foreign currency in domestic currency. Since \( P^* \) is assumed to be constant, the domestic price level \( P \) is expressed by the exchange rate \( \varepsilon \).

Now assume that a market agent’s utility depends on consumption and leisure so that:

\[ \alpha \log C + (1 - \alpha) \log (\bar{L} - L_t), \quad (2.58) \]

where \( \bar{L} - L_t \) denotes leisure, and \( \alpha (1 - \alpha) \) shows the portion of an individual’s limited budget that is devoted to consumption (leisure), with \( 0 \leq \alpha \leq 1 \).

Now let us assume that an individual’s leisure is an increasing function of real money holdings relative to consumption:

\[ \bar{L} - L_t = \bar{L} \left( \frac{M_t}{P_t} \right)^{\beta}, \quad (2.59) \]

with \( 0 < \beta < \frac{\alpha}{1 - \alpha} \). When combining equations (2.58) and (2.59), one yields:

\[ [\alpha - \beta(1 - \alpha)] \log C_t + \beta(1 - \alpha) \log \frac{M_t}{P_t}. \quad (2.60) \]
As Eq. (2.60) shows, real money holdings have now entered a market agent’s utility function and, as a result, provide an explanation why people may wish to hold money balances.

### 2.4.2 Cash-in-Advance Models of Money Demand

An alternative way of formulating the demand for money is to assume a cash-in-advance (CIA) restriction (Clower, 1967). The CIA model represents a rather extreme micro-founded transaction approach to the demand for money: money demand is determined by the need to make purchases rather than economizing. In a popular variant of the CIA model (Lucas, 1982), agents must acquire currency in period \( t - 1 \) to cover the consumption purchases they make in period \( t \) (Obstfeld & Rogoff, 1999, p. 547). The market agent’s challenge is to maximize the utility (\( U \)) of consumption (\( C \)) according to the following equation:

\[
U_t = \sum_{s=t}^{\infty} \beta^{s-t} u(C_s),
\]  

(2.61)

where \( \beta \) is the discount factor. Money does not enter the utility function. However, Eq. (2.61) is subject to the individual market agent’s budget constraint, which is:

\[
B_{t+1} + \frac{M_t}{P_t} = (1 + r)B_t + \frac{M_{t-1}}{P_t} + Y_t - C_t - T_t, \quad (2.62a)
\]

where \( B \) represents holdings of bonds issued by non-residents (we assume, for simplicity, that residents don’t issue debt), which are denominated in terms of real output, \( Y \) is real income and \( T \) are real taxes. The timing convention is that \( M_t \) is the nominal stock of money which is accumulated during period \( t \) and then carried over into \( t + 1 \).

What is more, the CIA constraint is:

\[
M_{t-1} \geq P_t C_t, \quad (2.62b)
\]

that is nominal consumption spending in \( t \) must be lower or equal the stock of nominal money held in \( t - 1 \). What about the interest rate (that is the opportunity cost) of money holdings? If the interest rate is positive, people wouldn’t want to hold money in excess of next period’s consumption spending, as they could earn a return by investing the money. If the nominal interest rate is zero, the CIA constraint is:

\[
M_{t-1} = P_t C_t. \quad (2.62c)
\]

Using (2.62c) to eliminate \( M_t \) and \( M_{t-1} \) from Eq. (2.62a) yields:

\[
B_{t+1} = (1 + r)B_t + Y_t - T_t - \frac{P_{t+1}}{P_t} + C_{t+1}; \quad (2.63)
\]
note that the last term on the right hand side of (2.63) is from substitution $M_t/P_t = (P_{t+1}/P_t)C_{t+1}$; its meaning will become clear shortly.

The intertemporal Euler condition can be calculated by maximizing the following function with respect to $B_s$:

$$U_t = \sum_{s=t}^{\infty} \beta^{s-t} u \left\{ \frac{P_{s-1}}{P_s} [ (1 + r) B_{s-1} - B_s + Y_{s-1} - T_{s-1}] \right\}. \quad (2.64)$$

Because we have $C_t = M_{t-1}/P_t$, with $M_{t-1}$ given in $t$, $C_t$ is predetermined – given by past history – in the individual’s maximization problem, and therefore not subject to choice on date $t$.

Differentiating (2.64) with respect to $B_s$ (for $s > t$) yields:

$$\frac{P_{s-1}}{P_s} u'(C_s) = (1 + r) \frac{P_s}{P_{s+1}} \beta u'(C_{s+1}). \quad (2.65)$$

Now we take recourse to the Fisher parity, which expresses the nominal interest rate as the real interest rate plus inflation:

$$1 + i_{s+1} = (1 + r)(P_{s+1}/P_s). \quad (2.66)$$

Dividing both sides of (2.65) by $1 + r$ and using (2.66) yields:

$$\frac{u'(C_s)}{1 + i_s} = (1 + r) \beta \frac{u'(C_{s+1})}{1 + i_{s+1}}. \quad (2.67)$$

According to Eq. (2.67) the additional cost for money held between $s - 1$ and $s$ is $i_s$. The cost is $i_{s+1}$ for money balances held between $s$ and $s + 1$. Note that consumption involves an additional cost, as the agent component must wait one full period between the date he converts bonds or output into cash and the date he can consume. In that sense the nominal interest rate acts as a consumption tax. In a stationary equilibrium (constant money growth), nominal interest rates and the implied consumption tax are constant, and Eq. (2.67) would equal the usual Euler equation (2.64).

It is important to note that the money demand equation for this kind of CIA model is determined by consumption, rather than by total income:

$$\frac{M_{t-1}}{P_t} = C_t. \quad (2.68)$$

In this simple CIA model, anticipated inflation does not affect money demand (as it would, for instance, in Friedman’s theory). As this is theoretically unappealing, Lucas and Stokey (1987) developed a model which combines the features of the CIA and the MUF approach. Cash needs to be held to finance consumption purchases, but the agent has some flexibility in allocating consumption between goods subject to the CIA constraint (cash goods) and goods that can be purchased in exchange
for securities (credit goods). This feature of the CIA constraint generates a non-zero interest rate elasticity of real money demand.

A simple variant of the CIA model as put forward by Helpman (1981) and Lucas (1982) gives consumers the opportunity to use cash acquired in period t for consumption later in period t. The producers of goods who receive the cash must hold it between periods. The CIA constraint (2.62b) is then:

$$M_t \geq P_t C_t.$$  \hspace{1cm} (2.69)

Here, inflation does not affect consumption decisions. However, the model holds that inflation influences the budget constraints of producers. The latter would hold money between periods in proportion to current sales, so that inflation becomes a production tax rather than a consumption tax in this model.\(^{10}\)

### 2.5 Estimating Money Demand Functions for the US and the Euro Area

#### 2.5.1 Money Demand in the US

During the 1980s, M2 became the primary intermediate target of US monetary policy. Since the early 1990s, however, the reliability of money measures as targets or indicators of monetary policy has been called into question. In 1993 the FOMC “downgraded” the role of M2. Though the “traditional” M2 relation broke down somewhere around 1990, there is some evidence that the disturbance might be a permanent upward shift in M2 velocity, which began in the early 1990s and was largely over by 1994. In terms of the monetary aggregates MZM and M2M, Carlson, Hoffman, Keen, and Rasche (2000) found strong evidence of a stable money demand relationships through the 1990s.

To revisit the issue, the long-run demand function for US monetary aggregates can be formulated as:

$$m_t - p_t = \beta_0 + \beta_1 y_t - \beta_2 i_{t, opportunity} + \beta_3 dum_t + \epsilon_t,$$  \hspace{1cm} (2.70)

where \(m_t\) is the stock of money, \(p_t\) is the GDP deflator, \(y_t\) is real output (all variables in natural logarithms). The opportunity costs of money holdings is the difference between the 3-months money market rate and the yield of money included in M2:

$$i_{t, opportunity} = \ln(1 + i_{t, 3-months}/100) - \ln(1 + i_{t, M2\ own\ rate}/100).$$  \hspace{1cm} (2.71)

Finally, \(dum\) is a dummy variable, taking the value of 0 from 1968-Q1 to 1989-Q4, then \(dum\) rises linearly to 1 until 1994-Q1 and remains at that level thereafter.

---

\(^{10}\)See in this context also the work of Aschauer and Greenwood (1983).
2.5 Estimating Money Demand Functions for the US and the Euro Area

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Long-run relation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_t$</td>
<td>-8.862</td>
<td>-0.872</td>
<td>-0.877</td>
<td>-0.862</td>
</tr>
<tr>
<td>(0.014)</td>
<td>(0.012)</td>
<td>(0.022)</td>
<td>(0.045)</td>
<td></td>
</tr>
<tr>
<td>$i^\text{opportunity}$</td>
<td>3.255</td>
<td>3.264</td>
<td>4.177</td>
<td>7.313</td>
</tr>
<tr>
<td>(0.296)</td>
<td>(0.275)</td>
<td>(0.485)</td>
<td>(0.905)</td>
<td></td>
</tr>
<tr>
<td>Dum</td>
<td>-0.139</td>
<td>0.139</td>
<td>0.115</td>
<td>0.089</td>
</tr>
<tr>
<td>(0.011)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.901</td>
<td>3.979</td>
<td>4.014</td>
<td>3.826</td>
</tr>
<tr>
<td>Unit root tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 2</td>
<td>-3.02 [0.03]</td>
<td>-3.42 [0.01]</td>
<td>-3.09 [0.02]</td>
<td>-2.89 [0.04]</td>
</tr>
<tr>
<td>Lag 4</td>
<td>-3.21 [0.02]</td>
<td>-3.63 [0.00]</td>
<td>-3.09 [0.02]</td>
<td>-2.95 [0.04]</td>
</tr>
<tr>
<td>Lag 6</td>
<td>-4.72 [0.00]</td>
<td>-5.25 [0.00]</td>
<td>-4.15 [0.00]</td>
<td>-3.79 [0.00]</td>
</tr>
<tr>
<td>II. ECM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lags</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$EC_{t-1}$</td>
<td>-0.334</td>
<td>-0.300</td>
<td>-0.158</td>
<td>-0.055</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.61</td>
<td>0.61</td>
<td>0.57</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Fig. 2.13 Long-run demand for US real M2, conventional

Source: Federal Reserve Bank of St. Louis, Thomson Financial; own calculations. Legend: The price level is approximated by the GDP deflator, $y_t$ is real output (in natural $ln$). The opportunity costs of money holdings is the difference between the 3-months money market rate and the yield of money included in M2: $i^\text{opportunity} = \ln(1 + i_{3\text{ mths}} / 100) - \ln(1 + i_{\text{own rate}} / 100); dum$ is a dummy variable, taking the value of 0 from 1970-Q4 to 1989-Q4, then dum rises linearly to 1 until 1994-Q4, and remains constant at that level thereafter. – $EC$ represents the error correction term of the first difference equation. – Lags in quarters. – (.) are standard errors, [.] $p$-values. – Results of the ADF-tests.

(Carlson, Hoffman, Keen, & Rasche, 2000). Figure 2.13 shows the cointegration results following the Johansen methodology for four sample periods.

The Johansen Procedure – an Overview

The finding that many macroeconomic time series may contain a unit root has encouraged the development of the theory of non-stationary time series analysis. Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. If such a stationary linear combination exists, the non-stationary time series are cointegrated. The stationary linear combination is called the cointegrating equation and may be interpreted as a long-run equilibrium relationship which exists among the variables under review.

The purpose of the cointegration test is to determine whether a group of non-stationary series are cointegrated or not. The presence of a cointegrating relation forms the basis of the vector error correction model (VEM). To
cointegration technique, we start with a vector autoregressive (VAR) model of order $p$:

$$y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + B x_t + \varepsilon_t$$  \hspace{1cm} \text{(2.72)}

where $y_t$ is a vector of non-stationary I(1) variables, $x_t$ is a $d$-vector of determinstic variables, and $\varepsilon_t$ is a vector of innovations. We can rewrite the VAR as:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + B x_t + \varepsilon_t,$$  \hspace{1cm} \text{(2.73)}

where

$$\Pi = \sum_{i=1}^{p} A_i - I$$  \hspace{1cm} \text{and}  \hspace{1cm} $$\Gamma_i = - \sum_{j=i+1}^{p} A_j.$$  \hspace{1cm} \text{(2.74)}

Granger’s representation theorem asserts that if the coefficient matrix $\Pi$ has reduced rank (such that $r < k$), then there exist $k \times r$ matrices $\alpha$ and $\beta$ each with rank $r$ such that $\Pi = \alpha \beta'$ and $\beta' y_t$ is I(0). $r$ is the number of cointegrating relations (the rank) and each column of $\beta$ is the cointegrating vector. The elements of $\alpha$ are the adjustment parameters in the VEC model. Johansen’s method is to estimate the $\Pi$ from an unrestricted VAR and to test whether we can reject the restrictions implied by the reduced rank of $\Pi$.

To determine the number of cointegrating relations conditional on the assumptions made about the trend, one can proceed sequentially from $r = 0$ to $r = k - 1$ until one fails to reject the null hypothesis. The trace statistic tests the null hypothesis of $r$ cointegrating relations against the alternative of $k$ cointegrating relations, where $k$ is the number of endogenous variables, for $r = 0, 1, \ldots, k - 1$. The alternative of $k$ cointegrating relations corresponds to the case where none of the series has a unit root and a stationary VAR may be specified in terms of the levels of all of the series. The trace statistic for the null hypothesis of $r$ cointegrating relations is computed as:

$$LR_{tr}(r|k) = - T \sum_{i=r+1}^{k} \log(1 - \lambda_i)$$  \hspace{1cm} \text{(2.75)}

where $\lambda_i$ is the $i$-th largest eigenvalue of the $\Pi$ matrix. The maximum eigenvalue statistic shows the results of testing the null hypothesis of $r$ cointegrating relations against the alternative of $r + 1$ cointegrating relations. This test statistic is computed as:
2.5 Estimating Money Demand Functions for the US and the Euro Area

\[ LR_{\text{max}}(r|r+1) = -T \log(1 + \lambda_{r+1}) = LR_{tr}(r|k) - LR_{tr}(r|r+1) \]  
\text{for } r = 0, 1, \ldots, k - 1. \tag{2.76}

The estimated coefficients of the long-run demand for real M2 exhibit economically plausible signs and magnitudes. For all sample periods, the income elasticity of money demand remains in a relatively narrow band of between 0.88 and 0.86. The variability of the interest rate elasticity is much stronger, though, varying between 3.26 for the period 1970-Q4 to 1989-Q4 and 7.31 for the total period under review. The results of the ADF-tests suggest that the null hypothesis of a unit root in the residuals can be rejected at standard levels for all estimations. That said, the findings seem to indicate that there is a long-run relation between real money demand, real output and the opportunity costs of money holdings in the US. Also, the error correction terms (EC) are statistically significant and have a negative sign, suggesting that deviations of real money holdings from the level suggested by the long-run equation are corrected over time.

Figure 2.14(a) displays the EC terms of the estimation results for the four sample periods. As can be seen, the EC terms fluctuate around the zero line, reflecting the fact that real M2 holdings tend to deviate temporarily from the long-run equilibrium value. What is more, the longer the sample periods are for which the coefficients were estimated, the higher is the volatility of the EC term. This could suggest that in the 1990s, in addition to the long-run determinants of money demand, numerous additional factors (shocks) might have affected money demand.

In a second step, the demand for real M2 balances is analysed using a different definition of the price level. In fact, the price level was calculated as a weighted average of the GDP deflator and house price index.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2_14.png}
\caption{Deviation of US real M2 holdings from equilibrium}
\end{figure}
\begin{flushright}
Source: Federal Reserve Bank of St. Louis, Thomson Financial; own calculations.
\end{flushright}
### Long-run relation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_t$</td>
<td>-0.779</td>
<td>-0.799</td>
<td>-0.806</td>
<td>-0.797</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>$i_t^{\text{opportunity}}$</td>
<td>3.522</td>
<td>3.458</td>
<td>3.767</td>
<td>4.292</td>
</tr>
<tr>
<td></td>
<td>(0.452)</td>
<td>(0.370)</td>
<td>(0.392)</td>
<td>(0.388)</td>
</tr>
<tr>
<td>$Dum$</td>
<td>–</td>
<td>0.141</td>
<td>0.131</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.016)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>$Constant$</td>
<td>4.505</td>
<td>4.676</td>
<td>4.732</td>
<td>4.646</td>
</tr>
</tbody>
</table>

Unit root tests:

<table>
<thead>
<tr>
<th>Lag 2</th>
<th>-3.96 [0.00]</th>
<th>-4.11 [0.00]</th>
<th>-3.98 [0.00]</th>
<th>-4.17 [0.00]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag 4</td>
<td>-3.43 [0.01]</td>
<td>-3.50 [0.01]</td>
<td>-3.41 [0.01]</td>
<td>-3.65 [0.00]</td>
</tr>
<tr>
<td>Lag 6</td>
<td>-4.59 [0.00]</td>
<td>-4.75 [0.00]</td>
<td>-4.63 [0.00]</td>
<td>-4.81 [0.00]</td>
</tr>
</tbody>
</table>

### ECM

<table>
<thead>
<tr>
<th>Lags</th>
<th>2</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$EC_{t-1}$</td>
<td>-0.193</td>
<td>-0.168</td>
<td>-0.146</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.52</td>
<td>0.54</td>
<td>0.54</td>
</tr>
</tbody>
</table>

---

**Fig. 2.15** Long-run demand for US real M2, amended

Source: Federal Reserve Bank of St. Louis, Thomson Financial; own calculations. **Legend:** The price level is the weighted average of the GDP deflator (80%) and the US housing price index (20%). $y_t$ is real output (in natural ln). The opportunity costs of money holdings is the difference between the 3-months money market rate and the yield of money included in M2: $i_t^{\text{opportunity}} = \ln(1 + i_{3\text{mths}}^t/100) - \ln(1 + i_{M2\text{own rate}}^t/100)$; dum is a dummy variable, taking the value of 0 from 1970-Q4 to 1989-Q4, the dum rises linearly to 1 until 1994-Q4, and remains constant at that level thereafter. $EC$ represents the error correction term of the first difference equation. – Lags in quarters. – (.) are standard errors, [.] p-values. – Results of the ADF-tests.

average of the GDP deflator and a US wide housing price index. The underlying idea is that market agents might hold nominal M2 balances in relation to a price level much broader defined than the GDP deflator (or, for that matter, the consumer price index). The estimation results are shown in Fig. 2.15.

As in the conventional estimation, income and interest elasticity of money demand have plausible signs and magnitudes. Especially in the period 1970-Q4 to 2005-Q4 the interest elasticity rises only slightly compared to the sub-periods and the conventional estimation results. Again, the $EC$ terms are statistically significant, suggesting that money holding disequilibria are reduced over time.

Figure 2.14(b) shows the EC terms for the estimates above. The volatility of the EC terms is much smaller than in the conventional estimates. With the parameters of the long-run demand function for real M2 being relatively stable for all sub-periods under consideration, it seems that using a weighted price level might be a promising approach to formulate and (re-)establish a stable demand for money function in the US.
2.5 Estimating Money Demand Functions for the US and the Euro Area

Stock Prices in the Demand for Money Function

To highlight the long-run relation between stock prices, dividend yields and central bank short-term interest rates, the arbitrage model of Cassola and Morana (2002) will be outlined. It can form a theoretical framework for integrating the stock market into the demand for money theory.

Money Demand

The demand for real money, $rm$, is defined as a function of output, $y$, and a vector of relative asset returns:

$$ rm_t = \phi_1 y_t - \phi' s_t, \quad (2.77) $$

where $s_t$ is a vector of yields of alternative assets and $\phi$ is a vector of parameters. It is assumed that there is a stationary combination of the yield spreads:

$$ \phi' s_t = \varepsilon_{rm,t} \sim I(0) \text{ where } \varepsilon_{rm,t} = \rho_1 \varepsilon_{rm,t-1} + \upsilon_t, \text{ is an AR}(1) $$

with $\rho_1 < 1$ and $\upsilon_t$ white noise.

Term structure of interest rates

The first arbitrage relationship is between short- and long-term yields:

$$ l_t = i_t + \phi_l + \varepsilon_{l,t}, \quad (2.78) $$

where $\varepsilon_{l,t} = \rho_2 \varepsilon_{l,t-1} + \upsilon_{l,t}$ with $\rho_2 < 1$ and $\upsilon_{l,t}$ white noise. The expectation theory of the term structure of interest rates can be written in logarithmic form as:

$$ l_t = \frac{1}{n} \sum_{j=0}^{n-1} E_t [i_{t+j}] + \phi_l, \quad (2.79) $$

where $l_t$ is expressed as an average of expected one period yields, $E_t [i_{t+j}]$, $\phi_l$ is a term premium and $n$ is the maturity of the bond. Substracting $i_t$ from both sides yields:

$$ l_t - i_t = \frac{1}{n} \sum_{j=0}^{n-1} E_t [i_{t+j} - i_t] + \phi_l \text{ with } \frac{1}{n} \sum_{j=0}^{n-1} E_t [i_{t+j} - i_t] = \varepsilon_{l,t}. \quad (2.80) $$
**Fisher Parity**

The second arbitrage relationship is the Fisher parity. It links the long-term asset return, $i_t$, with the short-term interest rate:

$$i_t = \phi_{fp} + \phi_2 \pi_t + \epsilon_{i,t}, \quad (2.81)$$

where $\phi_{fp}$ represents the real short-term interest rate plus the inflation risk premium and $\epsilon_{i,t} = \rho_3 \epsilon_{i,t-1} + \nu_{i,t}$ with $\rho_3 < 1$ and $\nu_{i,t}$ white noise.

The real asset return can be formulated as:

$$1 + \rho_t = (1 + \rho) \exp(\epsilon_{\rho,t}). \quad (2.82)$$

The Fisher parity for the short-term interest rate can be written as:

$$1 + R_t = (1 + \rho_{t+1})(1 + \pi_{t+1}^*) \exp(\epsilon_{\rho p,t+1}). \quad (2.83)$$

$R_t$ is the gross nominal return on a short-term investment for the period $t$ to $t+1$. $\epsilon_{\rho p,t} = \rho_4 \epsilon_{\rho p,t-1} + \nu_{\rho p,t}$ is the error term reflecting fluctuations in the real return and inflation with $\rho_4 < 1$ and $\nu_{\rho p,t}$ white noise.

The expected inflation is given by:

$$(1 + \pi_{t+1}^*) = (1 + \pi_t) \exp(\nu_{\pi^*,t}), \quad (2.84)$$

where $\nu_{\pi^*,t}$ is a stationary process.

Inserting (2.82) and (2.84) in (2.83) yields:

$$1 + R_t = (1 + \rho)(1 + \pi_{t+1}) \exp(\epsilon_{\rho p,t+1} + \epsilon_{\rho,t+1} + \nu_{\pi^*,t+1}). \quad (2.85)$$

For the long-term asset return, one yields the log-linear approximation:

$$i_t = \rho + \epsilon_{\rho p,t} + \pi_{t+1} + \epsilon_{\rho,t+1} + \nu_{\pi^*,t+1}, \quad (2.86)$$

which actually represents Eq. (2.81) if the unity coefficient on actual inflation is dropped. If (2.81) is inserted in (2.79) one yields:

$$l_t = \phi_{fp} + \phi_l + \phi_2 \pi_t + \epsilon_{i,t} + \epsilon_{i,t}, \quad (2.87)$$

which represents the long-run Fisher parity.
Stock Market and Output

The third arbitrage relation links the stock market to real output:

\[ f_t = \phi_f + \phi_3 y_t + \varepsilon_{f,t}, \]  

(2.88)

where \( f_t \) is the log real stock market capitalization and \( \varepsilon_{f,t} = \rho_5 \varepsilon_{f,t-1} + \nu_{f,t} \) with \( \rho_5 < 1 \) and \( \nu_{f,t} \) white noise. The present value model is:

\[ F_t = E_t \left[ \sum_{j=1}^{\infty} (1 + \nu)^{-j} D_{t+j} \right], \]  

(2.89)

where \( F_t \) is the real stock market capitalization, \( \nu \) is the real risk-adjusted discount rate and \( D_t \) is the real dividend paid at time \( t \). Assuming a constant rate of growth for dividends \( (g) \), the Gordon (1962) growth model is:

\[ F_t = \frac{1 + g}{\nu - g} D_t. \]  

(2.90)

If dividends are constant over time, the formula can be reduced to:

\[ \frac{D_t}{F_t} = \nu, \]  

(2.91)

where the dividend yield \( D_t/F_t \) equals the real risk-adjusted rate of return on capital.

The relation between real dividends and output is:

\[ D_t = k Y_t^{\phi_3} \exp(\varepsilon_{d,t}), \]  

(2.92)

with \( \varepsilon_{d,t} = \rho_6 \varepsilon_{d,t-1} + \nu_{d,t} \) with \( \rho_6 < 1 \) and \( \nu_{d,t} \) white noise, so that Eq. (2.88) can be written as:

\[ f_t = \ln \left( \frac{1 + g}{\nu - g} \right) + \ln k + \phi_3 y_t + \varepsilon_{d,t}. \]  

(2.93)

Equations (2.86) and (2.93) have important implications for the relation between the bond and stock markets. If dividends are assumed to be constant the equation above can be written as:

\[ f_t - \phi_3 y_t = -\ln \nu + \ln k + \varepsilon_{d,t}. \]  

(2.94)
The long-term real yield is:
\[ l_t - \phi_2 \pi_t = \rho + \varepsilon_{rp,t} + \phi l + \nu_{\pi^*,t} + \varepsilon_{l,t}. \] (2.95)

The relation between stock prices and the return on capital would be:
\[ -\ln \nu + \ln k + \varepsilon_{d,t} \approx \rho + \varepsilon_{rp,t} + \nu_{\pi^*,t} + \varepsilon_{l,t}. \] (2.96)

These relations suggest that there should be a stable linkage between the stock market, inflation, long- and short-term interest rates as set by the central bank.

Source: Cassola and Morana (2002).

2.5.2 Euro Area Money Demand 1980-Q1 to 2001-Q4

Empirical evidence supports the existence of a stable long-run relationship between money and output in the euro area (Bruggemann, Donati, & Warne, 2003). In fact, there is evidence that broad money demand has been more stable in the euro area than in other large economies (Calza & Sousa, 2003). First, some of the factors affecting money demand stability outside the euro area were country-specific. Second, the impact of financial innovation was relatively small in the euro area. Third, money demand in the euro area may have been more stable because it is an aggregation of money demand functions in individual countries. Figure 2.16 provides a selection of studies of the euro area demand function for broad money.

A first impression on the properties of the demand for money in the euro area can be derived from taking a look at the income velocity of M3 (Fig. 2.17(a)). In the period 1980-Q1 to the end of 2001 a linear function approximated the trend decline of M3 income velocity quite well. Thereafter, velocity appears to fall much stronger, indicating that money supply growth was markedly higher than nominal income growth. Figure 2.17(b) shows the deviations of actual income velocity from two trend lines, one calculated for 1980-Q1 to 2001-Q4 and one calculated for 1980-Q4 to 2005-Q4.

Given possible interdependencies among the variables which are typically seen as representing the explanatory variables of the long-run demand function for money, the latter cannot be estimated as a single equation. In fact, the demand for money should be estimated with a cointegration approach (Johansen, 1995). The individual equations of the system are as follows:

\[ m_t - p_t = \beta_{0,1} + \beta_{1,1} y_t + \beta_{2,1} i_t^g + \beta_{3,1} i_t^s + \beta_{4,1} i_t^p + \beta_{5,1} \pi_t + \varepsilon_{t,1} \] (2.97)

\[ i_t^s = \beta_{0,2} + \beta_{2,2} i_t^g + \beta_{4,2} i_t^p + \beta_{5,2} \pi_t + \varepsilon_{t,2} \] (2.98)
The first equation represents the demand function for real balances \((m - p_t)\), depending on real income \((y_t)\), the long- and short-rate (that is \(i^l_t\) and \(i^s_t\), respectively), the M3 own rate \((i^o_t)\), and inflation \((\pi_t)\), that is the first difference of the logarithm of the quarterly GDP deflator, annualized). The second equation relates the short-term interest rate to the long-term interest rate, the M3 own rate and inflation; combined the second equation captures the Fisher parity and the term structure of interest rates (Belke and Pollet, 2006). The time series are shown in Fig. 2.18.

Combining the variables in a vector \(X_t = (m - p_t, y_t, i^l_t, i^s_t, i^o_t, \pi_t)'\), the model can be formalized as \(\hat{\beta}'X_t = \varepsilon_t\), or:

\[
\hat{\beta}'X_t = \begin{bmatrix}
-1 & \beta_{1,1} & 0 & \beta_{3,1} & \beta_{4,2} & 0 \\
0 & 0 & \beta_{2,2} & -1 & \beta_{4,2} & \beta_{5,2}
\end{bmatrix}
\begin{bmatrix}
m_t - p_t \\
y_t \\
i^l_t \\
i^s_t \\
i^o_t \\
\pi_t
\end{bmatrix}
= \begin{bmatrix}
\varepsilon_{t,1} \\
\varepsilon_{t,2}
\end{bmatrix}
\]
### Fig. 2.17  Euro area M3 income, actual and trend, and deviations from trend
Source: ECB, Thomson Financial; own estimates. Income velocity was calculated by subtracting the stock of M3 from nominal GDP (all variables in logarithms). Period: 1980-Q1 to 2005-Q4.

### Fig. 2.18  Euro area money demand, data overview
For the period 1980-Q1 to 2001-Q4, the cointegration analysis between real stock of money, GDP, long- and short-term interest rates, the M3 own rate and inflation yields the following results:

\[
\hat{\beta}'X_t = \begin{bmatrix}
1 & -1.35 & 0 & 0.004 & 0 & 0 \\
0 & 0 & 0.28 & 1 & -1.71 & -0.81 \\
\end{bmatrix}
\begin{bmatrix}
\bar{m}_t - \bar{p}_t \\
y_t \\
i_t^l \\
i_t^s \\
i_t^o \\
\pi_t \\
\end{bmatrix}
\]

Chi-squared (4) = 0.68 [0.88]; standard errors in brackets.

Examining the first cointegration relation (that is the first row), which resembles a long-run money demand relation, one finds that the estimated coefficient on output is 1.35. It is greater than unity and of a similar magnitude indicated by earlier studies on euro area money demand (Brand & Cassola, 2000; Golinelli & Pastorello, 2002; Calza, Gartner, and Sousa, 2001). The demand on real money holdings increases (declines) if the short-term interest rate declines (rises), as suggested by the coefficient 0.004. Turning to the second vector (that is the second row), the coefficient on the long-term interest rate (0.28) suggests that long-term interest rates rise when short-term rates decline and vice versa. What is more, the yield on M3 holdings and inflation move up (down) when short-term rates fall (rise).

Overall, real money, real income, long- and short-term rates, the M3 own yield and inflation are trending together and form two long-run, or steady state, relations, where the first vector can be interpreted as the long-run money demand. Figure 2.19(a) and (b) show the residuals of the cointegration relations. The ADF-tests suggest that the null hypothesis of a unit root in the residuals can be rejected.

Alternatively, a money demand system might be modelled taking into account stock markets. In this case, a three equations system might be formulated:

\[
m_t - p_t = \beta_{0,1} + \beta_{1,1}y_t + \beta_{2,1}i_t^l + \beta_{3,1}i_t^s + \beta_{4,1}\pi_t + \beta_{5,1}s + \varepsilon_{t,1} \tag{2.99}
\]

\[
i_t^l = \beta_{0,2} + \beta_{3,2}i_t^s + \beta_{4,2}\pi_t + \varepsilon_{t,2} \tag{2.100}
\]

\[
s_t = \beta_{0,3} + \beta_{1,3}y_t + \varepsilon_{t,3}. \tag{2.101}
\]

The first equation represents the long-run demand function for real balances \((m_t - p_t)\), depending on real income \((y_t)\), long- and short-rates (that is \(i_t^l\) and \(i_t^s\), respectively), inflation \(\pi_t\) (first difference of quarterly changes in the GDP deflator, annualized) and real stock performance, \(s_t\). The second equation shows the combined Fisher parity and the term structure of interest rates. The third equation relates real stock market performance with real GDP. The result of the cointegration analysis for the period 1980-Q1 to 2001-Q4 is shown below:
134 Money and Credit Demand

(b) Residual of the combined test of the (i) term structure of interest rates and (ii) Fisher parity

Fig. 2.19 – Estimated cointegration relations for euro area M3 money demand over the period 1981-Q1 to 2001-Q4
Source: ECB; Thomson Financial; own calculations. – The ADF-tests reject the null hypothesis of a unit root in the residuals at the 5- and 1-percent level, respectively.

\[
\begin{bmatrix}
1 & -1.380 & 0.012 & 0 & 0 & 0 \\
0 & 0 & 1 & -0.197 & -0.266 & 0 \\
0 & 1 & 0 & 0 & 0 & -0.166
\end{bmatrix}
\]

\[
\begin{bmatrix}
m_t - p_t \\
y_t \\
\bar{i}_t \\
\bar{s}_t \\
\pi_t \\
s_t
\end{bmatrix}
\]

Chi-squared (5) = 6.04 [0.19]; standard errors in brackets.

The first vector might be interpreted as the long-run demand for real balances. The income elasticity is 1.38, while money demand responds negatively to rises in long-term interest rates. A long-run relation can be established between long- and short-term rates and inflation, as represented by the second vector. The third vector relates real output positively to real stock prices. The results of the difference equations – which estimate changes in real M3 holdings as a function of lagged changes in the other variables of the cointegration system (VAR in first differences) – are

<table>
<thead>
<tr>
<th>Lags</th>
<th>(\Delta(m_t - p_t))</th>
<th>(\Delta(y_t))</th>
<th>(\Delta(\bar{i}_t))</th>
<th>(\Delta(\bar{s}_t))</th>
<th>(\Delta(\pi_t))</th>
<th>(\Delta(s_t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(ECT1_{t-1})</td>
<td>-0.164 (3.74)</td>
<td>0.104 (1.90)</td>
<td>-1.944 (-0.39)</td>
<td>-21.15 (1.87)</td>
<td>19.93 (1.78)</td>
<td>0.710 (1.12)</td>
</tr>
<tr>
<td>(ECT2_{t-1})</td>
<td>-0.000 (-0.33)</td>
<td>0.001 (1.09)</td>
<td>-0.333 (-2.75)</td>
<td>0.112 (0.62)</td>
<td>0.310 (1.12)</td>
<td>-0.040 (-1.45)</td>
</tr>
<tr>
<td>(ECT3_{t-1})</td>
<td>0.035 (1.52)</td>
<td>-0.076 (-2.61)</td>
<td>5.626 (2.15)</td>
<td>5.918 (1.50)</td>
<td>-11.64 (-1.95)</td>
<td>0.639 (1.07)</td>
</tr>
<tr>
<td>R2</td>
<td>0.45</td>
<td>0.28</td>
<td>0.27</td>
<td>0.31</td>
<td>0.47</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Fig. 2.20 Results of the first difference estimates, 1980-Q1 to 2001-Q4; \(t\)-values in brackets
shown in Fig. 2.20. The error term of the long-run demand function of real M3, \( ECT_{t-1} \), has a negative sign, implying that excess real money holdings are reduced over time.

### 2.5.3 Euro Area Money Demand 1980-Q1 to 2006-Q1

Since 2001, euro area M3 growth has been constantly above the ECB’s 4\(\frac{1}{2}\) percent reference value. Hitherto stable standard money demand models fail to explain the observed monetary developments. One explanation for this phenomenon claims that an environment of increased macroeconomic uncertainty in conjunction with low asset yields has enhanced the preference for liquidity. Greiber and Lemke (2005) produced indicators (mainly based on financial market data) for measuring investor uncertainty. They show that uncertainty helps to explain the increase in euro area M3 over the period 2001 to 2004. In particular, a cointegrated money demand relationship can be established for samples that include these periods.

---

**Euro Area Money Demand Stability Reconsidered**

While a wide range of recent studies deals with the money demand relationship in the euro area, it has been analyzed almost exclusively on the basis of aggregate euro area data. Most of these studies exclusively use synthetic data for the pre-EMU period,\(^{11}\) but the more recent papers add data on the first years of EMU.\(^{12}\) Overviews are presented by Golinelli and Pastorello (2002) and Brand, Gerdesmeier & Roffia (2002).

Almost all papers find euro area money demand to be stable until the EMU started in 1999, even though they differ in many respects (sample, variables, estimation procedure, geographic area, aggregation method). A further outstanding result for studies with sample periods ending prior to 1999 is the higher stability of the area-wide compared to the country-specific money demand functions. It is, however, not clear how this can be explained properly; whether it is just a “statistical artefact” (Müller & Hahn, 2001), the positive influence of the traditionally stable German money demand (Calza & Sousa, 2003), or the neutralization of currency substitution movements across the union. On the other hand, as argued by Müller and Hahn (2001) and Hayo (1999), it is not clear whether the better stability properties of aggregate euro area money demand have persisted since the introduction of the euro.

---


The stability issue has received particular attention since M3 growth started to accelerate in 2001. Due to the strong M3 growth, Kontolemis (2002) finds evidence for money demand instability in the third quarter of 2001, the last observation in his sample. In a comprehensive stability analysis Bruggemann, Donati & Warne (2003) apply the fluctuation and Nyblom–type stability tests proposed by Hansen and Johansen (1999) and obtain mixed results but finally conclude that there are some specifications of long-run money demand that seem to be stable.

This result is challenged by Carstensen (2006) and Greiber and Lemke (2005). They argue that conventional money demand functions become unstable during the recent period of strong M3 growth and should be augmented with measures of macroeconomic or financial uncertainty, which account for the observation reported by the ECB (2003) that, following the terrorist attacks of September 2001 and the burst of the new economy bubble, large funds were reallocated into safe and liquid assets that are part of M3. However, these augmented specifications seem to be unable to explain the increase in M3 growth since the middle of 2004 (ECB, 2005; Alves, Marques, & Sousa, 2006). In contrast, Dreger and Wolters (2006) are still able to find a stable money demand function using data until the end of 2004. Therefore, the question whether long-run money demand is stable in the euro area still remains unsettled.

The focus on aggregate euro area data and euro area wide money demand stability may not be surprising since the ECB should be exclusively concerned with economic developments in the euro area as a whole. However, a disaggregate analysis on the basis of individual country data can lead to additional important insights both for EMU member countries and for the euro area as a whole (Belke & Gros, 2007; Belke & Heine, 2006). As concerns the individual EMU member countries and their central banks, they should be interested in the timely detection of national imbalances. Assuming that monetary aggregates and, in particular, money overhangs which are defined as the deviation of actual M3 from the money demand equilibrium, carry important information with respect to the state of the monetary and financial system, they should closely track the evolution of these quantities at the country level. This is ever more important if one follows Milton Friedman’s dictum that inflation is always and everywhere a monetary phenomenon because then money overhangs indicate future inflationary pressure for the respective country.

But a sensible measure of excess money is not necessarily invariant to the country of interest. This obviously holds for the 4.5% reference value that was derived by the ECB from aggregate developments in the euro area and, thus, disregards specific developments in the individual member countries. This may also hold true for more elaborate measures like the money overhang because the monetary and banking systems, the preferences of house-
holds and, hence, the money demand functions are probably not equal across countries. As concerns the euro area as a whole, there are at least three reasons why national developments should be of interest. First, for the optimal conduct of monetary policy it may prove beneficial to use national information if the national monetary transmission mechanisms are asymmetric (Belke & Heine, 2006; de Grauwe & Senegas, 2003). Second, and related to the first point, inflation forecasts constructed by aggregating country-specific models outperform inflation forecasts constructed by using aggregate euro area data only (Marcellino, Stock, & Watson, 2003). Similarly, country-specific inflation helps to explain area wide inflation even after controlling for aggregate macroeconomic information (Beck, Hubrich, & Marcellino, 2006; Carstensen, Hagen, Hossfeld, & Salazar Neaves, 2008). This implies that if monetary developments have predictive content for inflation, it should pay off to augment the aggregate information set with national money overhang measures. Third, even if the national variables did not carry additional information over aggregated variables, the construction of the ECB Governing Council would nevertheless entail considerable importance for national developments because the majority of the council members represent national central banks and may experience political pressure if the national developments diverge from the aggregate ones (Carstensen et al., 2008; Heinemann & Huefner, 2004). In such a situation, it is possible that they will feel committed to the countries they represent rather than to the euro area as a whole.

In this context, Carstensen et al. (2008) analyze the money demand functions of the four largest EMU countries and a four-country (EMU-4) aggregate. They identify reasonable and stable money demand relationships for Germany, France and Spain as well as the EMU-4 aggregate based on a M3 money aggregate. In the case of Italy, results are less clear. From the estimated money demand functions, they derive both EMU-4 and country-specific measures of the money overhang. They find that the EMU-4 M3 overhang measure strongly correlates with the country-specific measures, particularly since the start of EMU, and that these measures are useful for predicting country-specific inflation. The analyses show, however, that the aggregate money overhang is an important, but not an exhaustive indicator at the disaggregate level.

Similar results have recently been provided by Hamori and Hamori (2008). They analyzed the stability of the money demand function using panel data from January 1999 through March 2006, covering eleven euro area countries (Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Portugal, and Spain). The author, find that the money demand function was stable with respect to M3. This arguably supports the suitability of the ECB’s focus on M3 money supply in its monetary policy.
A cointegration system for real M3 holdings for the period 1980-Q1 to 2006-Q1, including the stock market and a measure of volatility of stock prices, might look as follows:

\[ m_t - p_t = \beta_{0,1} + \beta_{1,1}y_t + \beta_{2,1}i_t^l + \beta_{3,1}i_t^s + \beta_{4,1}s_t + \beta_{5,1}s_v t + \beta_{6,1}\pi_t + \varepsilon_{t,1} \]  
(2.102)

\[ i_t^s = \beta_{0,2} + \beta_{2,2}i_t^l + \beta_{6,2}\pi_t + \varepsilon_{t,2} \]  
(2.103)

\[ s_t = \beta_{0,3} + \beta_{1,3}y_t + \varepsilon_{t,3} \]  
(2.104)

The first equation represents the long-run demand function for real balances \((m_t - p_t)\), depending on real income \((y_t)\), long- and short interest rates (that is \(i_t^l\) and \(i_t^s\), respectively), real stock performance, \(s_t\), stock market volatility, \(s_v t\), and inflation \(\pi_t\). The second equation shows the combined Fisher parity and the term structure of interest rates. The third equation relates real stock market performance to real GDP. The real stock market performance and the volatility measure are shown in Fig. 2.21a, b.

The results of the cointegration analysis for the period 1980-Q1 to 2006-Q1 are given below:

\[
\begin{bmatrix}
1 & -1.251 & 0.232 & 0 & 0 & 0 & 0.008 \\
0 & 0 & 1 & -1.031 & -0.157 & -0.313 & 0 \\
0 & 1 & 0 & 0 & 0 & -0.154 & 0
\end{bmatrix}
\begin{bmatrix}
m_t - p_t \\
y_t \\
i_t^l \\
i_t^s \\
\pi_t \\
s_t \\
s_v t
\end{bmatrix}
\]

Chi-squared \((5) = 10.8 [0.05]\), standard errors in brackets.

(a) Real stock market performance\(^a\)

(b) Stock market volatility\(^b\)

Fig. 2.21 Euro area stock market measures
Source: ECB, Thomson Financial. \(^a\)Logarithm of the euro area stock market performance index minus the logarithm of GDP deflator. \(^b\)Moving standard deviation of weekly first differences of logarithms of the stock market performance index over a gliding 52-week window.
2.6 Credit Demand

When it comes to formulating models for credit demand, most studies include an economic activity variable (such as real GDP or industrial production) and financing costs (market interest rates or bank lending rates) as its main determinants. However, there seems to be no consensus in the literature about how economic activity affects credit demand. Some empirical findings point to a positive relation between the two variables based on the notion that economic growth would have a positive effect on expected income and profits. According to this argument, firms’ profit making enables private agents to support higher levels of indebtedness and, consequently, finance consumption and investments through credit (Kashyap, Stein, & Wilcox, 1993). An additional aspect would be that expectations of higher activity

13 In the following, we focus on the US. For credit demand analyses in the euro area see, for instance, De Nederlandsche Bank (2000) which estimates national equations for bank loans to the private sector in several EU countries, Japan and the USA. Vega (1989) focuses on aggregate credit demand in Spain. Also, see Calza, Gartner, and Sousa (2001) for an analysis of the demand for bank loans to the private sector in the euro area.
and productivity can lead to a larger number of projects becoming profitable which, in turn, entails a higher demand for credit to fund them.

By contrast, studies focusing on the US economy question the existence of a stable relationship between credit and economic activity. Some go even further and argue that, if any such relationship existed, it might actually turn out to be negative (Bernanke & Gertler, 1995; Friedman & Kuttner, 1993). The main line of argument is that an increase in contemporary productivity (as opposed to expected productivity) leads to a rise in output and, ultimately, profits. During expansionary phases, companies might prefer to rely more on internal sources of finance and reduce external financing. Similarly, households may want to take advantage of higher incomes in expansion phases to reduce their debt. On the other hand, in recessions, when both disposable income of households and firms’ profitability decline, households and corporations may increase their demand for bank credit in order to smooth out the impact of lower incomes and profits.

Most empirical credit demand studies include a measure of the cost of loans as an explanatory variable, and in many cases come up with a negative sign of its estimated coefficient. The negative relationship between the demand for loans and their cost appears to be rather uncontroversial, though some studies have pointed out that the price of loans should be adjusted to reflect the opportunity cost of bank loans (i.e. the cost of alternative sources of finance should be netted out as in Friedman & Kuttner, 1993). The underlying argument is that the demand for loans will depend not only on the rate of borrowed funds, but also on their relative price (that is relative to the cost of funds obtained from other internal or external sources). However, this issue is more relevant for non-financial corporations than for households since the latter have limited access to financing from sources other than the banking sector.

In the following, the demand function for bank loans in the US shall be analysed. Figure 2.23(a) shows the nominal growth of US bank loans and US GDP for the period 1969-Q1 to 2005-Q3. Eyeballing the time series suggests a positive relation

![Graph: (a) Bank loans and GDP growth, nominal (% y/y) (b) Income velocity of bank loans]

Fig. 2.23 US bank loans and economic activity
between economic expansion and loan supply. Figure 2.23(b) displays the income velocity of bank loans for the period 1968-Q1 to 2005-Q3, that is the difference between the logarithm of nominal GDP and the logarithm of the stock of bank loans. The series exhibits a downward trend, suggesting that, on average, the growth of nominal bank loans exceeded nominal output growth in the period under review.

The empirical model of the demand function for US bank loans shall be based on the following long-run relationship:

\[
l_t - p_t = \alpha + \beta_1 y_t + \beta_2 \text{def}_t + \beta_3 \Delta \text{cpi}_t + \beta_4 s_t + \beta_5 i^\text{prime}_t + \beta_6 i^5\text{-year}_t + \beta_7 i^\text{fftr}_t + \epsilon_t
\]

(2.105)

where \(l\), \(p\) and \(y\) stand for bank loans, GDP deflator and real GDP, respectively, in logarithms. Further, \(\text{def}\) stands for the default rate of US corporate bonds, and shall capture the degree of lenders’ risk aversion. \(\Delta \text{cpi}\) is the first difference in the natural logarithm of the consumer price index. \(s\) stands for the real stock market performance index (deflated with the consumer price index). \(i^\text{prime}_t\), \(i^5\text{-year}_t\) and \(i^\text{fftr}_t\) represent the prime loan rate, the 5-year bond rate and the Federal Funds Rate, respectively. Finally, \(\epsilon\) is the i.i.d. error term. Figure 2.24(a–d) shows the time series under review.

(a) Real bank loans and stock prices\(^a\)

(b) Interest rates

(c) Default rate\(^b\)

(d) Consumer price inflation\(^c\)

Fig. 2.24 Data overview

Source: Thomson Financial, Bloomberg, Federal Reserve Bank of St. Louis, Moody’s; own calculations. \(^a\)Bank loans deflated with the GDP deflator, stock prices with the consumer price index. \(^b\)US default rate, all US corporate bonds, Moody’s. \(^c\)First differences of natural logarithms.
The results of ADF-tests suggest that for all variables under review the null hypothesis of a unit root cannot be rejected. As a result, all variables are treated as non-stationary ($I(1)$). The cointegration properties of the data were analysed by using the Johansen methodology. For determining the optimal lag lengths of the cointegration system, a vector auto regression (VAR) model for $l - p$, def, $\Delta cpi$, $s$, $t \text{prime}_t$, $t \text{5-year}_t$ and $t \text{fr}_t$ was estimated for the period 1974-Q2 to 2005-Q3.

The criterion for selecting the optimal lag length of the VAR consists of choosing exactly that number of lags that is needed to eliminate the vector autocorrelation in the residuals. The optimal lag length was determined by calculating and comparing the empirical realisations of the information criteria for the VARs. The Akaike information criterion reaches its maximum at a lag of 8 quarters, the Schwarz-criterion at a lag of 1 and the Hannan-Quinn criterion at a lag of 2 quarters. A lag of 4 quarters suggested no autocorrelation in the residuals (Fig. 2.25).

The cointegration model allows for an unrestricted constant with a linear deterministic trend in the variables but not in the cointegration relationship. The test results in Fig. 2.26 reveal that one cannot reject the existence of one cointegration vector driving the time series at lag 4. However, the use of a system with 3 lags would strongly speak in favour of using at least 2 cointegration vectors.

Assuming two cointegration vectors, the long-run relation between the real stock of bank loans, output, interest rates, inflation and the stock market can be stated as follows (asymptotic standard errors in brackets):

<table>
<thead>
<tr>
<th>$H_0$: rank = $p$</th>
<th>Maximum eigenvalue test-statistic</th>
<th>95% critical value</th>
<th>Trace test-statistic</th>
<th>95% critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 0$</td>
<td>65.896*</td>
<td>52.363 [.001]</td>
<td>190.529*</td>
<td>159.529 [.000]</td>
</tr>
<tr>
<td>$p \leq 1$</td>
<td>41.891</td>
<td>46.231 [.136]</td>
<td>124.655</td>
<td>125.615 [.057]</td>
</tr>
<tr>
<td>$p \leq 2$</td>
<td>30.852</td>
<td>40.078 [.369]</td>
<td>82.764</td>
<td>95.753 [.278]</td>
</tr>
<tr>
<td>$p \leq 3$</td>
<td>21.756</td>
<td>33.877 [.627]</td>
<td>51.911</td>
<td>69.819 [.553]</td>
</tr>
</tbody>
</table>

*denotes the rejection of the null hypothesis at the .05 level. – Probabilities in brackets. The Trace-test and the Max-eigenvalue indicate 1 cointegrating eqn(s) at the .05 level.
2.6 Credit Demand

\[
\hat{\beta}'X_t = \begin{bmatrix}
1 & -1.683^{(0.09)} & 0 & 0.049^{(0.01)} & 0 & 0.089^{(0.02)} & -0.106^{(0.02)} \\
0 & 0 & 1.81^{(0.05)} & 1.653^{(0.33)} & -0.987^{(0.14)} & 0.149^{(0.18)} & 0.089^{(0.02)} \\
\end{bmatrix}
\]

Chi-squared (4) = 8.36 [0.079].

The first vector might be interpreted as the demand function for real bank loans. The income coefficient of 1.68 reflects the downward trend of the income velocity of bank loans. A possible explanation is that GDP might not capture the impact of wealth, which might also relevant to explain credit demand. What is more, real bank loan demand declines when borrowing costs rise. Credit demand rises if there is a rise in the 5-year interest rate, though.

The second vector captures four long-run relations: (i) the Fisher parity, linking long-term interest rates and inflation, (ii) the term structure of interest, establishing a long-run relation between short- and long-term interest rates, (iii) the relation between the lending rate and the riskless 5-year rate (a proxy for the credit spread) and (iv) the relation between real stock prices and bond yields.

Figure 2.27(a) shows the residuals of the long-run demand function of real bank loans. The ADF-test rejects the null hypothesis of a unit root. The ECT is scaled so that deviations from the long-term equilibrium relationship average zero over the sample period. If the ECT is above (below) the zero line, the level of real loans is above (below) the equilibrium level. In that sense, a positive (negative) ECT can be interpreted as an “over-supply” (under-supply) of real loans.

Figure 2.27(b) shows the ECT of the combined long-run relation between interest rates, inflation and real stock prices. The ADF-test rejects the null hypothesis of a

![Fig. 2.27](image-url) Estimated cointegration relations for a model of the demand for real bank loans in the US
unit root. Finally, it should be noted that the volatility of the ECT of the second cointegration vectors is – measured on the basis of the standard deviation – around six and a half times the volatility of the ECT of the first vector.

The use of a VECM provides the opportunity to specify the long- and short-run dynamics of the variables under review, while also capturing potential endogeneity of the determinants of credit demand. In particular, while the cointegrating vector is generally interpreted as a long-run equilibrium relationship, the estimates of the short-term dynamics help to characterise the process of adjustment towards the long-run equilibrium.

The coefficient of the error correction term of the long-run demand function for real bank loans (ECT1) is statistically significant for explaining changes in real loan demand (Fig. 2.28). The sign of the ECT1 is negative, suggesting that an oversupply of bank loans is corrected by a decline in real bank loans in future periods. The magnitude of the coefficient is rather small, though, suggesting that in case of deviations of real loans from their equilibrium level the adjustment towards the long-run level takes quite some time.

The results of the misspecification tests for the single difference equations are presented in Fig. 2.29. The null hypothesis of autocorrelation in the single equation residuals can be rejected for all equations. ARCH effect can be detected only for real stock prices. Except for $i_{5-year}$, no ARCH effect (at lag 4) can be detected. A violation of the normality assumption can be detected for $l_t - p_t$, $def_t$, $Δcpi$, and $s_t$. The multivariate test statistics for autocorrelation, normality and heteroscedasticity are satisfactory.

<table>
<thead>
<tr>
<th></th>
<th>$Δ(l_t-p_t)$</th>
<th>$Δ(p_{t-1})$</th>
<th>$Δ(def_t)$</th>
<th>$Δ(Δcpi)$</th>
<th>$Δ(i_{5-year})$</th>
<th>$Δ(i_{5-year}^{avg})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECT1</td>
<td>−0.053</td>
<td>−0.018</td>
<td>2.941</td>
<td>−0.028</td>
<td>−0.033</td>
<td>−3.518</td>
</tr>
<tr>
<td></td>
<td>−0.017</td>
<td>−0.013</td>
<td>−3.907</td>
<td>−0.422</td>
<td>−0.149</td>
<td>−1.243</td>
</tr>
<tr>
<td></td>
<td>[−3.08]</td>
<td>[−1.37]</td>
<td>[0.75]</td>
<td>[−0.06]</td>
<td>[−0.22]</td>
<td>[−2.83]</td>
</tr>
<tr>
<td>ECT2</td>
<td>0.003</td>
<td>0.001</td>
<td>−0.346</td>
<td>0.109</td>
<td>−0.025</td>
<td>0.497</td>
</tr>
<tr>
<td></td>
<td>−0.002</td>
<td>−0.002</td>
<td>−0.517</td>
<td>−0.056</td>
<td>−0.020</td>
<td>−0.164</td>
</tr>
<tr>
<td></td>
<td>[1.30]</td>
<td>[0.69]</td>
<td>[−0.66]</td>
<td>[1.95]</td>
<td>[−1.27]</td>
<td>[3.02]</td>
</tr>
<tr>
<td>R2</td>
<td>0.655</td>
<td>0.445</td>
<td>0.560</td>
<td>0.592</td>
<td>0.317</td>
<td>0.714</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.526</td>
<td>0.237</td>
<td>0.395</td>
<td>0.440</td>
<td>0.062</td>
<td>0.607</td>
</tr>
<tr>
<td>Sum sq. resid</td>
<td>0.007</td>
<td>0.004</td>
<td>374.3</td>
<td>4.364</td>
<td>0.545</td>
<td>37.87</td>
</tr>
<tr>
<td>S.E. equation</td>
<td>0.009</td>
<td>0.007</td>
<td>2.028</td>
<td>0.219</td>
<td>0.077</td>
<td>0.645</td>
</tr>
<tr>
<td>F-statistic</td>
<td>5.073</td>
<td>2.142</td>
<td>3.401</td>
<td>3.884</td>
<td>1.244</td>
<td>6.681</td>
</tr>
</tbody>
</table>

Determinant resid covariance (dof adj.) 2.47E−14
Determinant resid covariance 1.83E−15
Log likelihood 707.6439
Akaike information criterion −6.53403
Schwarz criterion 0.128981

Fig. 2.28  Estimated error correction terms of the VECM

Legend: Standard errors in (.) and t-statistics in [.].
Weak exogeneity tests can be performed on the equations for \(\text{def}, \Delta \text{cpi}, s, i_{t}^{\text{prime}}, i_{t}^{5\text{-year}}\) and \(i_{t}^{\text{fftr}}\) in order to determine whether, in the spirit of a general-to-specific approach, it would be legitimate to specify the demand for loans as a single equation model instead of a system. The test is performed by assessing the statistical significance of the coefficient of the ECTs in each of the equations of the system other than the equation for loans. If the ECT is found not to be significant in a specific equation, this implies that there is no information loss from excluding that equation from the system.

The tests show that the null hypothesis of weak exogeneity cannot be rejected for all equations except for the prime lending rate. This implies that the VECM approach cannot be reduced to a single equation. Although on the basis of the weak exogeneity it would be possible to exclude this variable from the system and proceed with a smaller VECM (conditioned on these variables), one may decide to continue to retain the full system. One implication of the test result is that, in order to describe the dynamics of the adjustment of real loans to its equilibrium level, it cannot be assumed that in case real loans deviate from equilibrium the return to it will necessarily be prompted only by adjustments in real loans themselves. In fact, deviations from equilibrium may lead to movements also in the prime lending rate.

Figure 2.30 shows impulse-response functions of the VECM. It shows the reaction of real bank loans to a Cholesky one standard deviation of (i) real GDP, (ii) the prime lending rate, (iii) the default rate and (iv) the Federal Funds Rate. An increase in real GDP tends to raise the demand for real bank loans until 7 quarters following...
The impulse. An increase in the prime lending rate reduces the demand for real bank loans; the effect remains negative throughout the period under review. A hike of the Federal Funds Rate, however, increases the demand for real bank loans in the first 9 quarters. Thereafter, the negative impact kicks in. Finally, an increase in the default rate reduces the demand for bank loans.

References


References


References


Monetary Economics in Globalised Financial Markets
Belke, A.; Polleit, T.
2009, XIII, 819 p. 282 illus., Hardcover
ISBN: 978-3-540-71002-8