



## Money Supply, Money Demand, and the Banking System

*There have been three great inventions since the beginning of time: fire, the wheel, and central banking.*

—Will Rogers

The supply and demand for money are crucial to many issues in macroeconomics. In Chapter 4, we discussed how economists use the term “money,” how the central bank controls the quantity of money, and how monetary policy affects prices and interest rates in the long run when prices are flexible. In Chapters 10 and 11, we saw that the money market is a key element of the *IS–LM* model, which describes the economy in the short run when prices are sticky. This chapter examines money supply and money demand more closely.

In Section 19-1 we see that the banking system plays a key role in determining the money supply. We discuss various policy instruments that the central bank can use to influence the banking system and alter the money supply. We also discuss some of the regulatory problems that central banks confront—an issue that rose in prominence during the financial crisis and economic downturn of 2008 and 2009.

In Section 19-2 we consider the motives behind money demand, and we analyze the individual household’s decision about how much money to hold. We also discuss how recent changes in the financial system have blurred the distinction between money and other assets and how this development complicates the conduct of monetary policy.

### 19-1 Money Supply

Chapter 4 introduced the concept of “money supply” in a highly simplified manner. In that chapter we defined the quantity of money as the number of dollars held by the public, and we assumed that the Federal Reserve controls the supply of money by increasing or decreasing the number of dollars in circulation through open-market operations. This explanation is a good starting point for understanding what determines the supply of money, but it is incomplete, because it omits the role of the banking system in this process. We now present a more complete explanation.

In this section we see that the money supply is determined not only by Fed policy but also by the behavior of households (which hold money) and banks (in which money is held). We begin by recalling that the money supply includes both currency in the hands of the public and deposits at banks that households can use on demand for transactions, such as checking account deposits. That is, letting  $M$  denote the money supply,  $C$  currency, and  $D$  demand deposits, we can write

$$\text{Money Supply} = \text{Currency} + \text{Demand Deposits}$$

$$M = C + D.$$

To understand the money supply, we must understand the interaction between currency and demand deposits and how Fed policy influences these two components of the money supply.

### 100-Percent-Reserve Banking

We begin by imagining a world without banks. In such a world, all money takes the form of currency, and the quantity of money is simply the amount of currency that the public holds. For this discussion, suppose that there is \$1,000 of currency in the economy.

Now introduce banks. At first, suppose that banks accept deposits but do not make loans. The only purpose of the banks is to provide a safe place for depositors to keep their money.

The deposits that banks have received but have not lent out are called **reserves**. Some reserves are held in the vaults of local banks throughout the country, but most are held at a central bank, such as the Federal Reserve. In our hypothetical economy, all deposits are held as reserves: banks simply accept deposits, place the money in reserve, and leave the money there until the depositor makes a withdrawal or writes a check against the balance. This system is called **100-percent-reserve banking**.

Suppose that households deposit the economy's entire \$1,000 in Firstbank. Firstbank's **balance sheet**—its accounting statement of assets and liabilities—looks like this:

**Firstbank's Balance Sheet**

Assets		Liabilities	
Reserves	\$1,000	Deposits	\$1,000

The bank's assets are the \$1,000 it holds as reserves; the bank's liabilities are the \$1,000 it owes to depositors. Unlike banks in our economy, this bank is not making loans, so it will not earn profit from its assets. The bank presumably charges depositors a small fee to cover its costs.

What is the money supply in this economy? Before the creation of Firstbank, the money supply was the \$1,000 of currency. After the creation of Firstbank, the money supply is the \$1,000 of demand deposits. A dollar deposited in a bank reduces currency by one dollar and raises deposits by one dollar, so the money supply remains the same. *If banks hold 100 percent of deposits in reserve, the banking system does not affect the supply of money.*

## Fractional-Reserve Banking

Now imagine that banks start to use some of their deposits to make loans—for example, to families who are buying houses or to firms that are investing in new plants and equipment. The advantage to banks is that they can charge interest on the loans. The banks must keep some reserves on hand so that reserves are available whenever depositors want to make withdrawals. But as long as the amount of new deposits approximately equals the amount of withdrawals, a bank need not keep all its deposits in reserve. Thus, bankers have an incentive to make loans. When they do so, we have **fractional-reserve banking**, a system under which banks keep only a fraction of their deposits in reserve.

Here is Firstbank's balance sheet after it makes a loan:

Assets		Liabilities	
Reserves	\$200	Deposits	\$1,000
Loans	\$800		

This balance sheet assumes that the *reserve–deposit ratio*—the fraction of deposits kept in reserve—is 20 percent. Firstbank keeps \$200 of the \$1,000 in deposits in reserve and lends out the remaining \$800.

Notice that Firstbank increases the supply of money by \$800 when it makes this loan. Before the loan is made, the money supply is \$1,000, equaling the deposits in Firstbank. After the loan is made, the money supply is \$1,800: the depositor still has a demand deposit of \$1,000, but now the borrower holds \$800 in currency. *Thus, in a system of fractional-reserve banking, banks create money.*

The creation of money does not stop with Firstbank. If the borrower deposits the \$800 in another bank (or if the borrower uses the \$800 to pay someone who then deposits it), the process of money creation continues. Here is the balance sheet of Secondbank:

Assets		Liabilities	
Reserves	\$160	Deposits	\$800
Loans	\$640		

Secondbank receives the \$800 in deposits, keeps 20 percent, or \$160, in reserve, and then loans out \$640. Thus, Secondbank creates \$640 of money. If this \$640 is eventually deposited in Thirdbank, this bank keeps 20 percent, or \$128, in reserve and loans out \$512, resulting in this balance sheet:

Assets		Liabilities	
Reserves	\$128	Deposits	\$640
Loans	\$512		

The process goes on and on. With each deposit and loan, more money is created.

Although this process of money creation can continue forever, it does not create an infinite amount of money. Letting  $rr$  denote the reserve–deposit ratio, the amount of money that the original \$1,000 creates is

$$\begin{aligned}
 \text{Original Deposit} &= \$1,000 \\
 \text{Firstbank Lending} &= (1 - rr) \times \$1,000 \\
 \text{Secondbank Lending} &= (1 - rr)^2 \times \$1,000 \\
 \text{Thirdbank Lending} &= (1 - rr)^3 \times \$1,000 \\
 \hline
 \text{Total Money Supply} &= [1 + (1 - rr) + (1 - rr)^2 \\
 &\quad + (1 - rr)^3 + \dots] \times \$1,000 \\
 &= (1/rr) \times \$1,000.
 \end{aligned}$$

Each \$1 of reserves generates  $\$(1/rr)$  of money. In our example,  $rr = 0.2$ , so the original \$1,000 generates \$5,000 of money.<sup>1</sup>

The banking system's ability to create money is the primary difference between banks and other financial institutions. As we first discussed in Chapter 3, financial markets have the important function of transferring the economy's resources from those households that wish to save some of their income for the future to those households and firms that wish to borrow to buy investment goods to be used in future production. The process of transferring funds from savers to borrowers is called **financial intermediation**. Many institutions in the economy act as financial intermediaries: the most prominent examples are the stock market, the bond market, and the banking system. Yet, of these financial institutions, only banks have the legal authority to create assets (such as checking accounts) that are part of the money supply. Therefore, banks are the only financial institutions that directly influence the money supply.

Note that although the system of fractional-reserve banking creates money, it does not create wealth. When a bank loans out some of its reserves, it gives borrowers the ability to make transactions and therefore increases the supply of money. The borrowers are also undertaking a debt obligation to the bank, however, so the loan does not make them wealthier. In other words, the creation of money by the banking system increases the economy's liquidity, not its wealth.

### A Model of the Money Supply

Now that we have seen how banks create money, let's examine in more detail what determines the money supply. Here we present a model of the money supply under fractional-reserve banking. The model has three exogenous variables:

<sup>1</sup> *Mathematical note:* The last step in the derivation of the total money supply uses the algebraic result for the sum of an infinite geometric series (which we used previously in computing the multiplier in Chapter 10). According to this result, if  $x$  is a number between  $-1$  and  $1$ , then

$$1 + x + x^2 + x^3 + \dots = 1/(1 - x).$$

In this application,  $x = (1 - rr)$ .

- The **monetary base**  $B$  is the total number of dollars held by the public as currency  $C$  and by the banks as reserves  $R$ . It is directly controlled by the Federal Reserve.
- The **reserve–deposit ratio**  $rr$  is the fraction of deposits that banks hold in reserve. It is determined by the business policies of banks and the laws regulating banks.
- The **currency–deposit ratio**  $cr$  is the amount of currency  $C$  people hold as a fraction of their holdings of demand deposits  $D$ . It reflects the preferences of households about the form of money they wish to hold.

Our model shows how the money supply depends on the monetary base, the reserve–deposit ratio, and the currency–deposit ratio. It allows us to examine how Fed policy and the choices of banks and households influence the money supply.

We begin with the definitions of the money supply and the monetary base:

$$M = C + D,$$

$$B = C + R.$$

The first equation states that the money supply is the sum of currency and demand deposits. The second equation states that the monetary base is the sum of currency and bank reserves. To solve for the money supply as a function of the three exogenous variables ( $B$ ,  $rr$ , and  $cr$ ), we first divide the first equation by the second to obtain

$$\frac{M}{B} = \frac{C + D}{C + R}.$$

Then divide both the top and bottom of the expression on the right by  $D$ .

$$\frac{M}{B} = \frac{C/D + 1}{C/D + R/D}.$$

Note that  $C/D$  is the currency–deposit ratio  $cr$ , and that  $R/D$  is the reserve–deposit ratio  $rr$ . Making these substitutions, and bringing the  $B$  from the left to the right side of the equation, we obtain

$$M = \frac{cr + 1}{cr + rr} \times B.$$

This equation shows how the money supply depends on the three exogenous variables.

We can now see that the money supply is proportional to the monetary base. The factor of proportionality,  $(cr + 1)/(cr + rr)$ , is denoted  $m$  and is called the **money multiplier**. We can write

$$M = m \times B.$$

Each dollar of the monetary base produces  $m$  dollars of money. Because the monetary base has a multiplied effect on the money supply, the monetary base is sometimes called **high-powered money**.

Here's a numerical example. Suppose that the monetary base  $B$  is \$800 billion, the reserve–deposit ratio  $rr$  is 0.1, and the currency–deposit ratio  $cr$  is 0.8. In this case, the money multiplier is

$$m = \frac{0.8 + 1}{0.8 + 0.1} = 2.0,$$

and the money supply is

$$M = 2.0 \times \$800 \text{ billion} = \$1,600 \text{ billion.}$$

Each dollar of the monetary base generates two dollars of money, so the total money supply is \$1,600 billion.

We can now see how changes in the three exogenous variables— $B$ ,  $rr$ , and  $cr$ —cause the money supply to change.

1. The money supply is proportional to the monetary base. Thus, an increase in the monetary base increases the money supply by the same percentage.
2. The lower the reserve–deposit ratio, the more loans banks make, and the more money banks create from every dollar of reserves. Thus, a decrease in the reserve–deposit ratio raises the money multiplier and the money supply.
3. The lower the currency–deposit ratio, the fewer dollars of the monetary base the public holds as currency, the more base dollars banks hold as reserves, and the more money banks can create. Thus, a decrease in the currency–deposit ratio raises the money multiplier and the money supply.

With this model in mind, we can discuss the ways in which the Fed influences the money supply.

### The Three Instruments of Monetary Policy

In previous chapters we made the simplifying assumption that the Federal Reserve controls the money supply directly. In fact, the Fed controls the money supply indirectly by altering either the monetary base or the reserve–deposit ratio. To do this, the Fed has at its disposal three instruments of monetary policy: open-market operations, reserve requirements, and the discount rate.

**Open-market operations** are the purchases and sales of government bonds by the Fed. When the Fed buys bonds from the public, the dollars it pays for the bonds increase the monetary base and thereby increase the money supply. When the Fed sells bonds to the public, the dollars it receives reduce the monetary base and thus decrease the money supply. Open-market operations are the policy instrument that the Fed uses most often. In fact, the Fed conducts open-market operations in New York bond markets almost every weekday.

**Reserve requirements** are Fed regulations that impose on banks a minimum reserve–deposit ratio. An increase in reserve requirements raises the reserve–deposit ratio and thus lowers the money multiplier and the money supply. Changes in reserve requirements are the least frequently used of the Fed's three policy instruments.

The **discount rate** is the interest rate that the Fed charges when it makes loans to banks. Banks borrow from the Fed when they find themselves with too few reserves to meet reserve requirements. The lower the discount rate, the cheaper are borrowed reserves, and the more banks borrow at the Fed's discount window. Hence, a reduction in the discount rate raises the monetary base and the money supply.

Although these three instruments—open-market operations, reserve requirements, and the discount rate—give the Fed substantial power to influence the money supply, the Fed cannot control the money supply perfectly. Bank discretion in conducting business can cause the money supply to change in ways the Fed did not anticipate. For example, banks may choose to hold **excess reserves**—that is, reserves above the reserve requirement. The higher the amount of excess reserves, the higher the reserve–deposit ratio, and the lower the money supply. As another example, the Fed cannot precisely control the amount banks borrow from the discount window. The less banks borrow, the smaller the monetary base, and the smaller the money supply. Hence, the money supply sometimes moves in ways the Fed does not intend.

### CASE STUDY

#### Bank Failures and the Money Supply in the 1930s

Between August 1929 and March 1933, the money supply fell 28 percent. As we discussed in Chapter 11, some economists believe that this large decline in the money supply was the primary cause of the Great Depression. But we did not discuss why the money supply fell so dramatically.

The three variables that determine the money supply—the monetary base, the reserve–deposit ratio, and the currency–deposit ratio—are shown in Table 19-1 for 1929 and 1933. You can see that the fall in the money supply cannot be attributed to a fall in the monetary base: in fact, the monetary base rose 18 percent over this period. Instead, the money supply fell because the money multiplier fell 38 percent. The money multiplier fell because the currency–deposit and reserve–deposit ratios both rose substantially.

Most economists attribute the fall in the money multiplier to the large number of bank failures in the early 1930s. From 1930 to 1933, more than 9,000 banks suspended operations, often defaulting on their depositors. The bank failures caused the money supply to fall by altering the behavior of both depositors and bankers.

Bank failures raised the currency–deposit ratio by reducing public confidence in the banking system. People feared that bank failures would continue, and they began to view currency as a more desirable form of money than demand deposits. When they withdrew their deposits, they drained the banks of reserves. The process of money creation reversed itself, as banks responded to lower reserves by reducing their outstanding balance of loans.

In addition, the bank failures raised the reserve–deposit ratio by making bankers more cautious. Having just observed many bank runs, bankers became apprehensive about operating with a small amount of reserves. They therefore

**TABLE 19-1****The Money Supply and Its Determinants: 1929 and 1933**

	August 1929	March 1933
<b>Money Supply</b>	<b>26.5</b>	<b>19.0</b>
Currency	3.9	5.5
Demand deposits	22.6	13.5
<b>Monetary Base</b>	<b>7.1</b>	<b>8.4</b>
Currency	3.9	5.5
Reserves	3.2	2.9
<b>Money Multiplier</b>	<b>3.7</b>	<b>2.3</b>
Reserve–deposit ratio	0.14	0.21
Currency–deposit ratio	0.17	0.41

Source: Adapted from Milton Friedman and Anna Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, N.J.: Princeton University Press, 1963), Appendix A.

increased their holdings of reserves to well above the legal minimum. Just as households responded to the banking crisis by holding more currency relative to deposits, bankers responded by holding more reserves relative to loans. Together these changes caused a large fall in the money multiplier.

Although it is easy to explain why the money supply fell, it is more difficult to decide whether to blame the Federal Reserve. One might argue that the monetary base did not fall, so the Fed should not be blamed. Critics of Fed policy during this period make two counterarguments. First, they claim that the Fed should have taken a more vigorous role in preventing bank failures by acting as a *lender of last resort* when banks needed cash during bank runs. This would have helped maintain confidence in the banking system and prevented the large fall in the money multiplier. Second, they point out that the Fed could have responded to the fall in the money multiplier by increasing the monetary base even more than it did. Either of these actions would likely have prevented such a large fall in the money supply, which in turn might have reduced the severity of the Great Depression.

Since the 1930s, many policies have been put into place that make such a large and sudden fall in the money multiplier less likely today. Most important, the system of federal deposit insurance protects depositors when a bank fails. This policy is designed to maintain public confidence in the banking system and thus prevents large swings in the currency–deposit ratio. Deposit insurance has a cost: in the late 1980s and early 1990s, for example, the federal government incurred the large expense of bailing out many insolvent savings-and-loan institutions. Yet deposit insurance helps stabilize the banking system and the money supply. That is why, during the financial crisis of 2008–2009, the Federal Deposit Insurance Corporation raised the amount guaranteed from \$100,000 to \$250,000 per depositor. ■



## Bank Capital, Leverage, and Capital Requirements

The model of the banking system presented in this chapter is simplified. That is not necessarily a problem: after all, all models are simplified. But it is worth drawing attention to one particular simplifying assumption.

In the bank balance sheets presented so far, a bank takes in deposits and uses those deposits to make loans or to hold reserves. Based on this discussion, you might think that it does not take any resources to open up a bank. That is, however, not true. Starting a bank requires some capital. That is, the bank owners must start with some financial resources to get the business going. Those resources are called **bank capital** or, equivalently, the equity of the bank's owners.

Here is what a more realistic balance sheet for a bank would look like:

**A Bank's Balance Sheet**

Assets		Liabilities and Owners' Equity	
Reserves	\$200	Deposits	\$750
Loans	\$500	Debt	\$200
Securities	\$300	Capital (owners' equity)	\$50

The bank obtains resources from its owners, who provide capital, and also by taking in deposits and issuing debt. It uses these resources in three ways. Some funds are held as reserves; some are used to make bank loans; and some are used to buy financial securities, such as government or corporate bonds. The bank allocates its resources among these asset classes, taking into account the risk and return that each offers and any regulations that restrict its choices. The reserves, loans, and securities on the left side of the balance sheet must equal, in total, the deposits, debt, and capital on the right side of the balance sheet.

This business strategy relies on a phenomenon called **leverage**, which is the use of borrowed money to supplement existing funds for purposes of investment. The *leverage ratio* is the ratio of the bank's total assets (the left side of the balance sheet) to bank capital (the one item on the right side of the balance sheet that represents the owners' equity). In this example, the leverage ratio is  $\$1000/\$50$ , or 20. This means that for every dollar of capital that the bank owners have contributed, the bank has \$20 of assets and, thus, \$19 of deposits and debts.

One implication of leverage is that, in bad times, a bank can lose much of its capital very quickly. To see how, let's continue with this numerical example. If the bank's assets fall in value by a mere 5 percent, then the \$1,000 of assets are now worth only \$950. Because the depositors and debt holders have the legal right to be paid first, the value of the owners' equity falls to zero. That is, when the leverage ratio is 20, a 5-percent fall in the value of the bank assets leads to a 100-percent fall in bank capital. The fear that bank capital may be running out, and thus that depositors may not be fully repaid, is typically what generates bank runs when there is no deposit insurance.

One of the restrictions that bank regulators put on banks is that the banks must hold sufficient capital. The goal of such a **capital requirement** is to ensure

that banks will be able to pay off their depositors. The amount of capital required depends on the kind of assets a bank holds. If the bank holds safe assets such as government bonds, regulators require less capital than if the bank holds risky assets such as loans to borrowers whose credit is of dubious quality.

In 2008 and 2009 many banks found themselves with too little capital after they had incurred losses on mortgage loans and mortgage-backed securities. The shortage of bank capital reduced bank lending, contributing to a severe economic downturn. (This event was discussed in a Case Study in Chapter 11.) In response to this problem, the U.S. Treasury, working together with the Federal Reserve, started putting public funds into the banking system, increasing the amount of bank capital and making the U.S. taxpayer a part owner of many banks. The goal of this unusual policy was to recapitalize the banking system so bank lending could return to a more normal level.

## 19-2 Money Demand

We now turn to the other side of the money market and examine what determines money demand. In previous chapters, we used simple money demand functions. We started with the quantity theory, which assumes that the demand for real balances is proportional to income. That is, the quantity theory assumes

$$(M/P)^d = kY,$$

where  $k$  is a constant measuring how much money people want to hold for every dollar of income. We then considered a more general and realistic money demand function that assumes the demand for real money balances depends on both the interest rate and income:

$$(M/P)^d = L(i, Y).$$

We used this money demand function when we discussed the link between money and prices in Chapter 4 and when we developed the *IS–LM* model in Chapters 10 and 11.

There is, of course, much more to say about what determines how much money people choose to hold. Just as studies of the consumption function rely on microeconomic models of the consumption decision, studies of the money demand function rely on microeconomic models of the money demand decision. In this section we first discuss in broad terms the different ways to model money demand. We then develop one prominent model.

Recall that money serves three functions: it is a unit of account, a store of value, and a medium of exchange. The first function—money as a unit of account—does not by itself generate any demand for money, because one can quote prices in dollars without holding any. By contrast, money can serve its other two functions only if people hold it. Theories of money demand emphasize the role of money either as a store of value or as a medium of exchange.

## Portfolio Theories of Money Demand

Theories of money demand that emphasize the role of money as a store of value are called **portfolio theories**. According to these theories, people hold money as part of their portfolio of assets. The key insight is that money offers a different combination of risk and return than other assets. In particular, money offers a safe (nominal) return, whereas the prices of stocks and bonds may rise or fall. Thus, some economists have suggested that households choose to hold money as part of their optimal portfolio.<sup>2</sup>

Portfolio theories predict that the demand for money should depend on the risk and return offered by money and by the various assets households can hold instead of money. In addition, money demand should depend on total wealth, because wealth measures the size of the portfolio to be allocated among money and the alternative assets. For example, we might write the money demand function as

$$(M/P)^d = L(r_s, r_b, E\pi, W),$$

where  $r_s$  is the expected real return on stock,  $r_b$  is the expected real return on bonds,  $E\pi$  is the expected inflation rate, and  $W$  is real wealth. An increase in  $r_s$  or  $r_b$  reduces money demand, because other assets become more attractive. An increase in  $E\pi$  also reduces money demand, because money becomes less attractive. (Recall that  $-E\pi$  is the expected real return to holding money.) An increase in  $W$  raises money demand, because greater wealth means a larger portfolio.

From the standpoint of portfolio theories, we can view our money demand function,  $L(i, Y)$ , as a useful simplification. First, it uses real income  $Y$  as a proxy for real wealth  $W$ . Second, the only return variable it includes is the nominal interest rate, which is the sum of the real return on bonds and expected inflation (that is,  $i = r_b + E\pi$ ). According to portfolio theories, however, the money demand function should include the expected returns on other assets as well.

Are portfolio theories useful for studying money demand? The answer depends on which measure of money we are considering. The narrowest measures of money, such as  $M1$ , include only currency and deposits in checking accounts. These forms of money earn zero or very low rates of interest. There are other assets—such as savings accounts, Treasury bills, certificates of deposit, and money market mutual funds—that earn higher rates of interest and have the same risk characteristics as currency and checking accounts. Economists say that money ( $M1$ ) is a **dominated asset**: as a store of value, it exists alongside other assets that are always better. Thus, it is not optimal for people to hold money as part of their portfolio, and portfolio theories cannot explain the demand for these dominated forms of money.

Portfolio theories are more plausible as theories of money demand if we adopt a broad measure of money. The broad measures include many of those assets that

<sup>2</sup> James Tobin, "Liquidity Preference as Behavior Toward Risk," *Review of Economic Studies* 25 (February 1958): 65–86.

dominate currency and checking accounts.  $M2$ , for example, includes savings accounts and money market mutual funds. When we examine why people hold assets in the form of  $M2$ , rather than bonds or stock, the portfolio considerations of risk and return may be paramount. Hence, although the portfolio approach to money demand may not be plausible when applied to  $M1$ , it may be a good theory to explain the demand for  $M2$ .

### CASE STUDY

## Currency and the Underground Economy

How much currency are you holding right now in your wallet? How many \$100 bills?

In the United States today, the amount of currency per person is about \$3,000. About half of that is in \$100 bills. Most people find this fact surprising, because they hold much smaller amounts and in smaller denominations.

Some of this currency is used by people in the underground economy—that is, by those engaged in illegal activity such as the drug trade and by those trying to hide income to evade taxes. People whose wealth was earned illegally may have fewer options for investing their portfolio, because by holding wealth in banks, bonds, or stock, they assume a greater risk of detection. For criminals, currency may not be a dominated asset: it may be the best store of value available.

Some economists point to the large amount of currency in the underground economy as one reason that some inflation may be desirable. Recall that inflation is a tax on the holders of money, because inflation erodes the real value of money. A drug dealer holding \$20,000 in cash pays an inflation tax of \$2,000 per year when the inflation rate is 10 percent. The inflation tax is one of the few taxes those in the underground economy cannot evade.<sup>3</sup> ■

## Transactions Theories of Money Demand

Theories of money demand that emphasize the role of money as a medium of exchange are called **transactions theories**. These theories acknowledge that money is a dominated asset and stress that people hold money, unlike other assets, to make purchases. These theories best explain why people hold narrow measures of money, such as currency and checking accounts, as opposed to holding assets that dominate them, such as savings accounts or Treasury bills.

Transactions theories of money demand take many forms, depending on how one models the process of obtaining money and making transactions. All these theories assume that money has the cost of earning a low rate of return and the benefit of making transactions more convenient. People decide how much money to hold by trading off these costs and benefits.

<sup>3</sup> To read more about the large quantity of currency, see Case M. Sprenkle, “The Case of the Missing Currency,” *Journal of Economic Perspectives* 7 (Fall 1993): 175–184.

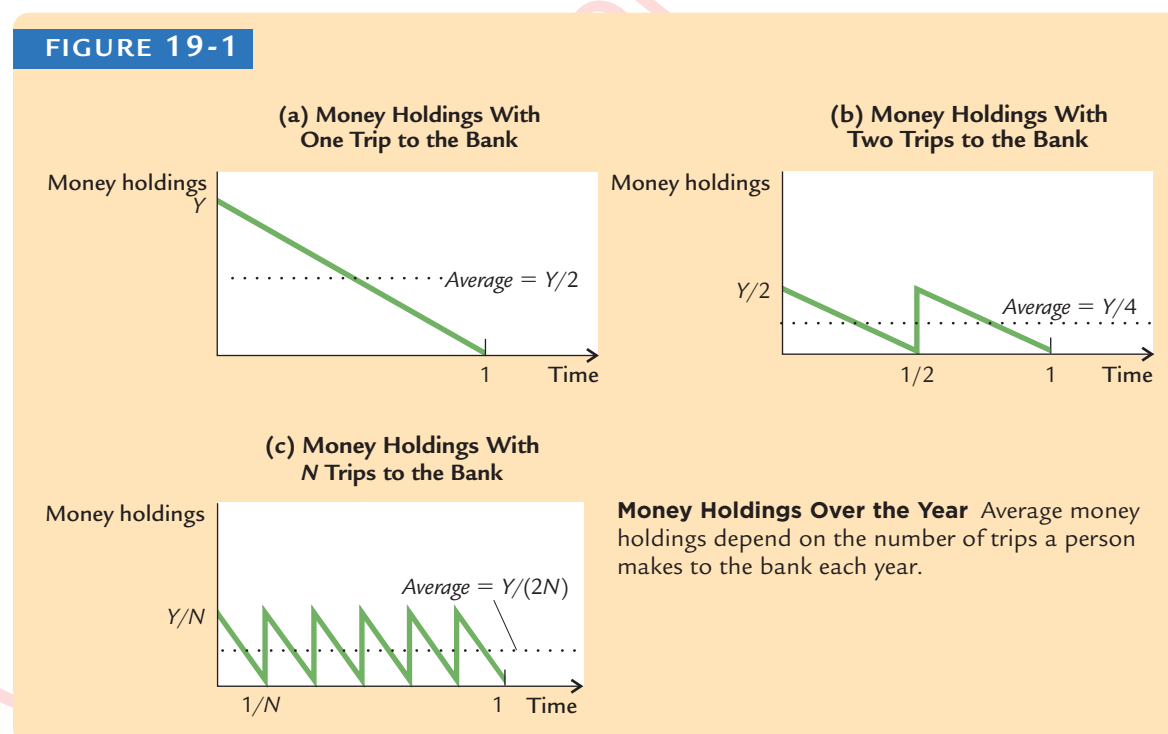
To see how transactions theories explain the money demand function, let's develop one prominent model of this type. The **Baumol–Tobin model** was developed in the 1950s by economists William Baumol and James Tobin, and it remains a leading theory of money demand.<sup>4</sup>

### The Baumol–Tobin Model of Cash Management

The Baumol–Tobin model analyzes the costs and benefits of holding money. The benefit of holding money is convenience: people hold money to avoid making a trip to the bank every time they wish to buy something. The cost of this convenience is the forgone interest they would have received had they left the money deposited in a savings account that paid interest.

To see how people trade off these benefits and costs, consider a person who plans to spend  $Y$  dollars gradually over the course of a year. (For simplicity, assume that the price level is constant, so real spending is constant over the year.) How much money should he hold in the process of spending this amount? That is, what is the optimal size of average cash balances?

Consider the possibilities. He could withdraw the  $Y$  dollars at the beginning of the year and gradually spend the money. Panel (a) of Figure 19-1 shows his



<sup>4</sup> William Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* 66 (November 1952): 545–556; and James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," *Review of Economics and Statistics* (August 1956): 241–247.

money holdings over the course of the year under this plan. His money holdings begin the year at  $Y$  and end the year at zero, averaging  $Y/2$  over the year.

A second possible plan is to make two trips to the bank. In this case, he withdraws  $Y/2$  dollars at the beginning of the year, gradually spends this amount over the first half of the year, and then makes another trip to withdraw  $Y/2$  for the second half of the year. Panel (b) of Figure 19-1 shows that money holdings over the year vary between  $Y/2$  and zero, averaging  $Y/4$ . This plan has the advantage that less money is held on average, so the individual forgoes less interest, but it has the disadvantage of requiring two trips to the bank rather than one.

More generally, suppose the individual makes  $N$  trips to the bank over the course of the year. On each trip, he withdraws  $Y/N$  dollars; he then spends the money gradually over the following  $1/N$ th of the year. Panel (c) of Figure 19-1 shows that money holdings vary between  $Y/N$  and zero, averaging  $Y/(2N)$ .

The question is, what is the optimal choice of  $N$ ? The greater  $N$  is, the less money the individual holds on average and the less interest he forgoes. But as  $N$  increases, so does the inconvenience of making frequent trips to the bank.

Suppose that the cost of going to the bank is some fixed amount  $F$ . We can view  $F$  as representing the value of the time spent traveling to and from the bank and waiting in line to make the withdrawal. For example, if a trip to the bank takes 15 minutes and a person's wage is \$12 per hour, then  $F$  is \$3. Also, let  $i$  denote the interest rate; because money does not bear interest,  $i$  measures the opportunity cost of holding money.

Now we can analyze the optimal choice of  $N$ , which determines money demand. For any  $N$ , the average amount of money held is  $Y/(2N)$ , so the forgone interest is  $iY/(2N)$ . Because  $F$  is the cost per trip to the bank, the total cost of making trips to the bank is  $FN$ . The total cost the individual bears is the sum of the forgone interest and the cost of trips to the bank:

$$\begin{aligned}\text{Total Cost} &= \text{Forgone Interest} + \text{Cost of Trips} \\ &= iY/(2N) + FN.\end{aligned}$$

The larger the number of trips  $N$ , the smaller the forgone interest, and the larger the cost of going to the bank.

Figure 19-2 shows how total cost depends on  $N$ . There is one value of  $N$  that minimizes total cost. The optimal value of  $N$ , denoted  $N^*$ , is<sup>5</sup>

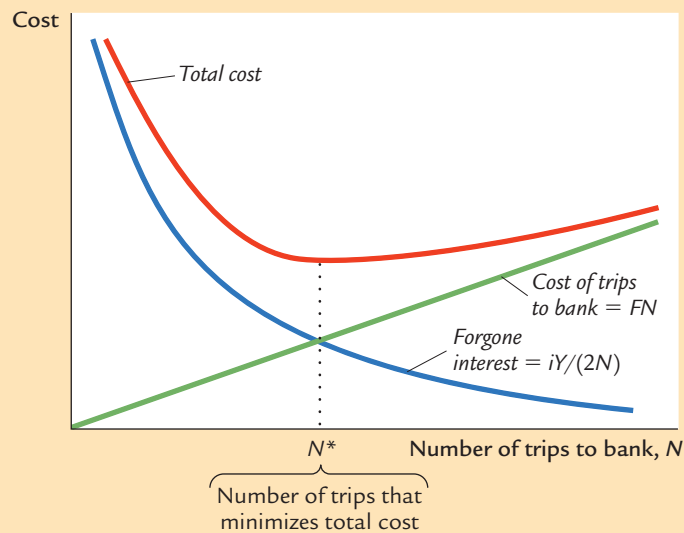
$$N^* = \sqrt{\frac{iY}{2F}}.$$

<sup>5</sup> *Mathematical note:* Deriving this expression for the optimal choice of  $N$  requires simple calculus. Differentiate total cost  $C$  with respect to  $N$  to obtain

$$dC/dN = -iYN^{-2}/2 + F.$$

At the optimum,  $dC/dN = 0$ , which yields the formula for  $N^*$ .

FIGURE 19-2

**The Cost of Money Holding**

Forgone interest, the cost of trips to the bank, and total cost depend on the number of trips  $N$ . One value of  $N$ , denoted  $N^*$ , minimizes total cost.

Average money holding is

$$\begin{aligned} \text{Average Money Holding} &= Y/(2N^*) \\ &= \sqrt{\frac{YF}{2i}}. \end{aligned}$$

This expression shows that the individual holds more money if the fixed cost of going to the bank  $F$  is higher, if expenditure  $Y$  is higher, or if the interest rate  $i$  is lower.

So far, we have been interpreting the Baumol–Tobin model as a model of the demand for currency. That is, we have used it to explain the amount of money held outside of banks. Yet one can interpret the model more broadly. Imagine a person who holds a portfolio of monetary assets (currency and checking accounts) and nonmonetary assets (stocks and bonds). Monetary assets are used for transactions but offer a low rate of return. Let  $i$  be the difference in the return between monetary and nonmonetary assets, and let  $F$  be the cost of transforming nonmonetary assets into monetary assets, such as a brokerage fee. The decision about how often to pay the brokerage fee is analogous to the decision about how often to make a trip to the bank. Therefore, the Baumol–Tobin model describes this person's demand for monetary assets. By showing that money demand depends positively on expenditure  $Y$  and negatively on the interest rate  $i$ , the model provides a microeconomic justification for the money demand function,  $L(i, Y)$ , that we have used throughout this book.

One implication of the Baumol–Tobin model is that any change in the fixed cost of going to the bank  $F$  alters the money demand function—that is, it

changes the quantity of money demanded for any given interest rate and income. It is easy to imagine events that might influence this fixed cost. The spread of automatic teller machines, for instance, reduces  $F$  by reducing the time it takes to withdraw money. Similarly, the introduction of Internet banking reduces  $F$  by making it easier to transfer funds among accounts. On the other hand, an increase in real wages increases  $F$  by increasing the value of time. And an increase in banking fees increases  $F$  directly. Thus, although the Baumol–Tobin model gives us a very specific money demand function, it does not give us reason to believe that this function will necessarily be stable over time.

### CASE STUDY

## Empirical Studies of Money Demand

Many economists have studied the data on money, income, and interest rates to learn more about the money demand function. One purpose of these studies is to estimate how money demand responds to changes in income and the interest rate. The sensitivity of money demand to these two variables determines the slope of the  $LM$  curve; it thus influences how monetary and fiscal policy affect the economy.

Another purpose of the empirical studies is to test the theories of money demand. The Baumol–Tobin model, for example, makes precise predictions for how income and interest rates influence money demand. The model's square-root formula implies that the income elasticity of money demand is  $1/2$ : a 10-percent increase in income should lead to a 5-percent increase in the demand for real balances. It also says that the interest elasticity of money demand is  $1/2$ : a 10-percent increase in the interest rate (say, from 10 percent to 11 percent) should lead to a 5-percent decrease in the demand for real balances.

Most empirical studies of money demand do not confirm these predictions. They find that the income elasticity of money demand is larger than  $1/2$  and that the interest elasticity is smaller than  $1/2$ . Thus, although the Baumol–Tobin model may capture part of the story behind the money demand function, it is not completely correct.

One possible explanation for the failure of the Baumol–Tobin model is that some people may have less discretion over their money holdings than the model assumes. For example, consider a person who must go to the bank once a week to deposit her paycheck; while at the bank, she takes advantage of her visit to withdraw the currency needed for the coming week. For this person, the number of trips to the bank,  $N$ , does not respond to changes in expenditure or the interest rate. Because  $N$  is fixed, average money holdings [which equals  $Y/(2N)$ ] are proportional to expenditure and insensitive to the interest rate.

Now imagine that the world is populated with two sorts of people. Some obey the Baumol–Tobin model, so they have income and interest elasticities of  $1/2$ . The others have a fixed  $N$ , so they have an income elasticity of 1 and an interest elasticity of zero. In this case, the overall demand for money looks like a



weighted average of the demands of the two groups. The income elasticity will be between  $1/2$  and  $1$ , and the interest elasticity will be between  $1/2$  and zero, as the empirical studies find.<sup>6</sup> ■

### Financial Innovation, Near Money, and the Demise of the Monetary Aggregates

Traditional macroeconomic analysis groups assets into two categories: those used as a medium of exchange as well as a store of value (currency, checking accounts) and those used only as a store of value (stocks, bonds, savings accounts). The first category of assets is called “money.” In this chapter we have discussed its supply and demand.

Although the distinction between monetary and nonmonetary assets remains a useful theoretical tool, in recent years it has become more difficult to use in practice. In part because of the deregulation of banks and other financial institutions, and in part because of improved computer technology, the past two decades have seen rapid financial innovation. Monetary assets such as checking accounts once paid no interest; today they earn market interest rates and are comparable to nonmonetary assets as stores of value. Nonmonetary assets such as stocks and bonds were once inconvenient to buy and sell; today mutual funds allow depositors to hold stocks and bonds and to make withdrawals simply by writing checks from their accounts. These nonmonetary assets that have acquired some of the liquidity of money are called **near money**.

The existence of near money complicates monetary policy by making the demand for money unstable. Because money and near money are close substitutes, households can easily switch their assets from one form to the other. Such changes can occur for minor reasons and do not necessarily reflect changes in spending. Thus, the velocity of money becomes unstable, and the quantity of money gives faulty signals about aggregate demand.

One response to this problem is to use a broad definition of money that includes near money. Yet, because there is a continuum of assets in the world with varying characteristics, it is not clear how to choose a subset to label “money.” Moreover, if we adopt a broad definition of money, the Fed’s ability to control this quantity may be limited, because many forms of near money have no reserve requirement.

The instability in money demand caused by near money has been an important practical problem for the Federal Reserve. In February 1993, Fed Chairman Alan Greenspan announced that the Fed would pay less attention to the monetary aggregates than it had in the past. The aggregates, he said, “do not appear to be giving reliable indications of economic developments and price pressures.” It’s easy to see why he reached this conclusion when he did. Over the preceding

---

<sup>6</sup> To learn more about the empirical studies of money demand, see Stephen M. Goldfeld and Daniel E. Sichel, “The Demand for Money,” *Handbook of Monetary Economics*, vol. 1 (Amsterdam: North-Holland, 1990), 299–356; and David Laidler, *The Demand for Money: Theories and Evidence*, 3rd ed. (New York: Harper & Row, 1985).

12 months,  $M1$  had grown at an extremely high 12-percent rate, while  $M2$  had grown at an extremely low 0.5-percent rate. Depending on how much weight was given to each of these two measures, monetary policy was either very loose, very tight, or somewhere in between.

Since then, the Fed has conducted policy by setting a target for the *federal funds rate*, the short-term interest rate at which banks make loans to one another. It adjusts the target interest rate in response to changing economic conditions. Under such a policy, the money supply becomes endogenous: it is allowed to adjust to whatever level is necessary to keep the interest rate on target. Chapter 14 presented a dynamic model of aggregate demand and aggregate supply in which an interest rate rule for the central bank is explicitly incorporated into the analysis of short-run economic fluctuations.

### 19-3 Conclusion

Money is at the heart of much macroeconomic analysis. Models of money supply and money demand can help shed light on the long-run determinants of the price level and the short-run causes of economic fluctuations. The rise of near money in recent years has shown that there is still much to be learned. Building reliable microeconomic models of money and near money remains a central challenge for macroeconomists.

### Summary

1. The system of fractional-reserve banking creates money, because each dollar of reserves generates many dollars of demand deposits.
2. The supply of money depends on the monetary base, the reserve–deposit ratio, and the currency–deposit ratio. An increase in the monetary base leads to a proportionate increase in the money supply. A decrease in the reserve–deposit ratio or in the currency–deposit ratio increases the money multiplier and thus the money supply.
3. The Federal Reserve changes the money supply using three policy instruments. It can increase the monetary base by making an open-market purchase of bonds or by lowering the discount rate. It can reduce the reserve–deposit ratio by relaxing reserve requirements.
4. To start a bank, the owners must contribute some of their own financial resources, which become the bank’s capital. Because banks are highly leveraged, however, a small decline in the value of their assets can potentially have a major impact on the value of bank capital. Bank regulators require that banks hold sufficient capital to ensure that depositors can be repaid.

5. Portfolio theories of money demand stress the role of money as a store of value. They predict that the demand for money depends on the risk and return on money and alternative assets.
6. Transactions theories of money demand, such as the Baumol–Tobin model, stress the role of money as a medium of exchange. They predict that the demand for money depends positively on expenditure and negatively on the interest rate.
7. Financial innovation has led to the creation of assets with many of the attributes of money. These near moneys make the demand for money less stable, which complicates the conduct of monetary policy.

### KEY CONCEPTS

Reserves	Money multiplier	Capital requirement
100-percent-reserve banking	High-powered money	Portfolio theories
Balance sheet	Open-market operations	Dominated asset
Fractional-reserve banking	Reserve requirements	Transactions theories
Financial intermediation	Discount rate	Baumol–Tobin model
Monetary base	Excess reserves	Near money
Reserve–deposit ratio	Bank capital	
Currency–deposit ratio	Leverage	

### QUESTIONS FOR REVIEW

1. Explain how banks create money.
2. What are the three ways in which the Federal Reserve can influence the money supply?
3. Why might a banking crisis lead to a fall in the money supply?
4. Explain the difference between portfolio and transactions theories of money demand.
5. According to the Baumol–Tobin model, what determines how often people go to the bank? What does this decision have to do with money demand?
6. In what way does the existence of near money complicate the conduct of monetary policy? How has the Federal Reserve responded to this complication?

### PROBLEMS AND APPLICATIONS

1. The money supply fell from 1929 to 1933 because both the currency–deposit ratio and the reserve–deposit ratio increased. Use the model of the money supply and the data in Table 19-1 to answer the following hypothetical questions about this episode.
  - a. What would have happened to the money supply if the currency–deposit ratio had risen but the reserve–deposit ratio had remained the same?

- b. What would have happened to the money supply if the reserve–deposit ratio had risen but the currency–deposit ratio had remained the same?
- c. Which of the two changes was more responsible for the fall in the money supply?
2. To increase tax revenue, the U.S. government in 1932 imposed a two-cent tax on checks written on deposits in bank accounts. (In today's dollars, this tax was about 25 cents per check.)
- a. How do you think the check tax affected the currency–deposit ratio? Explain.
- b. Use the model of the money supply under fractional-reserve banking to discuss how this tax affected the money supply.
- c. Now use the *IS–LM* model to discuss the impact of this tax on the economy. Was the check tax a good policy to implement in the middle of the Great Depression?
3. Give an example of a bank balance sheet with a leverage ratio of 10. If the value of the bank's assets rises by 5 percent, what happens to the value of the owners' equity in this bank? How large a decline in the value of bank assets would it take to reduce this bank's capital to zero?
4. Suppose that an epidemic of street crime sweeps the country, making it more likely that your wallet will be stolen. Using the Baumol–Tobin model, explain (in words, not equations) how this crime wave will affect the optimal frequency of trips to the bank and the demand for money.
5. Let's see what the Baumol–Tobin model says about how often you should go to the bank to withdraw cash.
- a. How much do you buy per year with currency (as opposed to checks or credit cards)? This is your value of  $Y$ .
- b. How long does it take you to go to the bank? What is your hourly wage? Use these two figures to compute your value of  $F$ .
- c. What interest rate do you earn on the money you leave in your bank account? This is your value of  $i$ . (Be sure to write  $i$  in decimal form—that is, 6 percent should be expressed 0.06.)
- d. According to the Baumol–Tobin model, how many times should you go to the bank each year, and how much should you withdraw each time?
- e. In practice, how often do you go to the bank, and how much do you withdraw?
- f. Compare the predictions of the Baumol–Tobin model to your behavior. Does the model describe how you actually behave? If not, why not? How would you change the model to make it a better description of your behavior?
6. In Chapter 4, we defined the velocity of money as the ratio of nominal expenditure to the quantity of money. Let's now use the Baumol–Tobin model to examine what determines velocity.
- a. Recalling that average money holdings equal  $Y/(2N)$ , write velocity as a function of the number of trips to the bank  $N$ . Explain your result.
- b. Use the formula for the optimal number of trips to express velocity as a function of expenditure  $Y$ , the interest rate  $i$ , and the cost of a trip to the bank  $F$ .
- c. What happens to velocity when the interest rate rises? Explain.
- d. What happens to velocity when the price level rises? Explain.
- e. As the economy grows, what should happen to the velocity of money? (*Hint*: Think about how economic growth will influence  $Y$  and  $F$ .)
- f. Suppose now that the number of trips to the bank is fixed rather than discretionary. What does this assumption imply about velocity?