Understanding Monetary Policy Implementation

Huberto M. Ennis
Research Department
Federal Reserve Bank of Richmond
Huberto.Ennis@rich.frb.org

Todd Keister
Research and Statistics Group
Federal Reserve Bank of New York
Todd.Keister@ny.frb.org

May 5, 2008

Abstract

The Federal Reserve implements its monetary policy objectives by intervening in the interbank market for reserves. In particular, it aims to change the supply of reserves available to commercial banks so that the (average) interest rate in this market equals an announced target rate. A recent change in legislation will give the Federal Reserve greater flexibility in this process by allowing it to pay interest on reserve balances. A combination of this change and recent events in financial markets has renewed interest in the process of monetary policy implementation. This article presents a simple analytical framework for understanding this process. We use the framework to illustrate the main factors that influence a central bank's ability to keep the market interest rate close to a target level, and we discuss how paying interest on reserves can be a useful policy tool in this regard.

This paper was written for the Federal Reserve Bank of Richmond's *Economic Quarterly*. Some of the material resulted from our participation in the Federal Reserve System task force created to study paying interest on reserves. We are very grateful to the other members of this group, who patiently taught us many of the things that we discuss here. We also would like to thank Kevin Bryan, Yash Mehra, John Walter, and John Weinberg for useful comments on a previous draft. All remaining errors are, of course, our own. The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of New York, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

1 Introduction

Over the last two decades, central banks around the world have adopted a common approach to monetary policy that involves targeting the value of a short-term interest rate. In the United States, for example, the Federal Open Market Committee (FOMC) announces a rate that it wishes to prevail in the federal funds market, where commercial banks lend balances held at the Federal Reserve to each other on an overnight basis. Changes in this short-term interest rate eventually translate into changes in other interest rates in the economy and thereby influence the overall level of prices and of real economic activity.

Once a target interest rate is announced, the problem of implementation arises: How can a central bank ensure that the relevant market interest rate stays at or near the chosen target? The Federal Reserve has a variety of tools available to influence the behavior of the interest rate in the federal funds market (called the *fed funds rate*). In general, the Fed aims to adjust the total supply of reserve balances so that it equals demand at exactly the target rate of interest. This process necessarily involves some estimation, since the Fed does not know the exact demand for reserve balances, nor does it completely control the supply in the market.

A critical issue in the implementation process, therefore, is the sensitivity of the market interest rate to unanticipated changes in supply and/or demand. If small estimation errors lead to large swings in the interest rate, a central bank will find it difficult to effectively implement monetary policy, that is, to consistently hit the target rate. The degree of sensitivity depends on a variety of factors related to the design of the implementation process, such as the time period over which banks are required to hold reserves and what interest rate, if any, a central bank pays on reserve balances.

The ability to consistently hit a target interest rate plays a critical role in a central bank's communication policy. The overall effectiveness of monetary policy depends, in part, on individuals' perceptions of the central bank's actions and objectives. If the market interest rate were to deviate consistently from the central bank's announced target, individuals might question whether these deviations simply represent glitches in the implementation process or whether they instead represent an unannounced change in the stance of monetary policy. Sustained deviations of the average fed funds rate from the FOMC's target in August 2007, for example, led some media commentators to claim that the Fed had engaged in a "stealth easing," taking actions that lowered the market interest rate without announcing a change in the official target.[1] In such times, the ability to consistently hit a target interest rate allows the central bank to clearly (and credibly) communicate its policy to market participants.

Under most circumstances, the Fed changes the total supply of reserve balances available to commercial banks by exchanging government bonds or other securities for reserves in an open market operation. Occasionally, the Fed also provides reserves directly to certain banks through its discount window. In some situations, the Fed has developed other, ad hoc methods of influencing the supply and distribution of reserves in the market.

For example, during the recent period of financial turmoil, the market's ability to smoothly distribute reserves across banks became partially impaired, which led to significant fluctuations in the average fed funds rate both during the day and across days. In December 2007, partly to address these problems the Fed introduced the Term Auction Facility (TAF), a bimonthly auction of a fixed quantity of reserve balances to all banks eligible to borrow at the discount window. In

[1] See, for example, "A 'Stealth Easing' by the Fed?" *Business Week*, August 17, 2007.

principle, the TAF has increased these banks' ability to directly access reserves and, in this way, has helped ease the pressure on the market to redistribute reserves and avoid abnormal fluctuations in the market rate. Such operations, of course, need to be managed so as to achieve the ultimate goal of implementing the chosen target interest rate. Balancing the demand and supply of reserves is at the very core of this problem.

This article presents a simple analytical framework for understanding the process of monetary policy implementation and the factors that influence a central bank's ability to keep the market interest rate close to a target level. We present this framework graphically, focusing on how different features of the implementation process affect the sensitivity of the market interest rate to unanticipated changes in supply or demand. We discuss the current approach used by the Fed, including the use of reserve maintenance periods to decrease this sensitivity. We also show how this framework can be used to study a wide range of issues related to monetary policy implementation. The mathematical model behind our graphical analysis is presented in the Appendix.

In 2006, the U.S. Congress enacted legislation that will give the Fed authority to pay interest on reserve balances beginning in October, 2011. We use our simple framework to illustrate how the ability to pay interest on reserves can be a useful policy tool for a central bank. In particular, we show how paying interest on reserves can decrease the sensitivity of the market interest rate to estimation errors and thus enable a central bank to better achieve its desired interest rate.

The model we present uses the basic approach to reserve management introduced by Poole (1968) and subsequently advanced by many others (see, for example, Dotsey 1991, Guthrie and Wright 2000, Clouse and Dow 2002, and Bartolini, Bertola, and Prati 2002). The specific details of our formalization closely follow those in Ennis and Weinberg (2007), after some additional simplifications. We conduct all of our analysis graphically. Ennis and Weinberg (2007) focused on the interplay between daylight credit and the Fed's overnight treatment of bank reserves. In this article, we take a more comprehensive view of the process of monetary policy implementation and we investigate several important topics, such as the role of reserve maintenance periods, which were left unexplored by Ennis and Weinberg (2007).

2 U.S. monetary policy implementation

Banks hold reserve balances in accounts at the Federal Reserve in order to satisfy reserve requirements and to be able to make interbank payments. During the day, banks can also access funds by obtaining an overdraft from their reserve accounts at the Fed. The terms by which the Fed provides daylight credit are one of the factors determining the demand for reserves by banks.

To adjust their reserve holdings, banks can borrow and lend balances in the fed funds market, which operates daily from 9:30 AM to 6:30 PM. A bank wanting to decrease its reserve holdings, for example, can do so in this market by making unsecured, overnight loans to other banks.

The fed funds market plays a crucial role in monetary policy implementation because this is where the Federal Reserve intervenes to pursue its policy objectives. The stance of monetary policy is decided by the FOMC, which selects a target for the overnight interest rate prevailing in this market. The Committee then instructs the Open Market Desk to adjust, via open market operations, the supply of reserve balances so as to steer the market interest rate toward the selected target.[2]

[2] See Hilton and Hrung (2007) for a more detailed overview of the Fed's monetary policy implementation procedures.

The Desk conducts open market operations largely by arranging repurchase agreements (repos) with primary dealers in a sealed-bid, discriminatory price auction. Repos involve selling reserve balances in exchange for bonds with the explicit agreement that the transaction will be reversed at maturity. Repos usually have overnight maturity, but the Desk also employs other maturities (for example, two-day and two-week repos are commonly used). Open market operations are typically conducted early in the morning when the market for repos is most active.

In these open market operations, the Desk tries to set the supply of reserve balances as close as possible to the level that would drive the market-clearing interest rate to equal the target rate. An essential step in this process is accurately forecasting both aggregate reserve demand and those changes in the existing supply of reserve balances that are due to *autonomous factors* beyond the Fed's control, such as payments into and out of the Treasury's account and changes in the quantity of currency in circulation. Forecasting errors will lead the actual supply of reserve balances to deviate from the intended level and, hence, will cause the market rate to diverge from the target rate even if reserve demand is perfectly anticipated.

Reserve requirements in the U.S. are calculated as a proportion of the quantity of transaction deposits on a bank's balance sheet during a two-week period prior to the start of the maintenance period. These requirements can be met through a combination of vault cash and reserve balances held at the Fed. During the two-week reserve maintenance period, a bank's end-of-day reserve balances must, on average, equal the reserve requirement minus the quantity of vault cash held during the previous computation period. Reserve requirements make a large portion of the demand for reserve balances fairly predictable, which simplifies monetary policy implementation.

Reserve maintenance periods allow banks to spread their reserve holdings over time without having to scramble for funds in order to meet a requirement at the end of each day. However, near the end of the maintenance period this averaging effect tends to lose force. On the last day of the period, a bank has a level of remaining requirement that must be met on that day. This generates a fairly inelastic demand for reserve balances and makes implementing a target interest rate more challenging. For this reason, the Fed allows banks holding excess or deficient balances at the end of a maintenance period to carry over those balances and use them in satisfying up to 4 percent of next period requirement.

If a bank finds itself short of reserves at the end of the maintenance period, even after taking into account the carryover possibilities, it has several options. It can try to find a counterparty late in the day offering an acceptable interest rate. However, this may not be feasible because of an aggregate shortage of reserve balances or because of the existence of trading frictions in this market. A second alternative is to borrow at the discount window of its corresponding Federal Reserve Bank.[3] The discount window offers collateralized overnight loans of reserves to banks that have previously pledged appropriate collateral. Discount window loans are typically charged an interest rate that is 100 basis points above the target Federal Funds rate, although changing the size of this gap is possible and has been used, at times, as a policy instrument. Finally, if the bank does not have the appropriate collateral, or chooses not to borrow at the discount window for other reasons, it will be charged a penalty fee proportional to the amount of the shortage.

Currently, banks earn no interest on the reserve balances they hold in their accounts at the Federal Reserve. This situation may soon change: The Financial Services Regulatory Relief Act of 2006 allows the Fed to begin paying interest on reserve balances in October, 2011. The Act also includes provisions that give the Fed more flexibility in determining reserve requirements,

[3] There are twelve regions and corresponding Reserve banks in the Federal Reserve System. For each commercial bank, the corresponding Reserve bank is that of the region where the commercial bank is headquartered.

including the ability to eliminate the requirements altogether. This legislation thus opens the door to potentially substantial changes in the way the Fed implements monetary policy. To evaluate the best approach within the new, broader set of alternatives, it seems useful to develop a simple analytical framework that is able to address many of the relevant aspects of the problem. We introduce and discuss such a framework in the sections that follow.

3 The Demand for Reserves

In this section, we present a simple framework that is useful for understanding banks' demand for reserves.[4] In this framework, a bank holds reserves primarily to satisfy reserve requirements, although other factors, such as the desire to make interbank payments, may also play a role. Since banks cannot fully predict the timing of payments, they face uncertainty about the net outflows from their reserve accounts and, therefore, are typically not able to exactly satisfy their reserve requirement. Instead, they must balance the possibility of holding excess reserve balances – and the associated opportunity cost – against the possibility of being penalized for a reserve deficiency. A bank's demand for reserves results from optimally balancing these two concerns.

3.1 The basic framework

We assume banks are risk neutral and maximize expected profits. At the beginning of the day, banks can borrow and lend reserves in a competitive interbank market. Let R be the quantity of reserves chosen by a bank in the interbank market. The central bank affects the supply of reserves in this market by conducting open market operations. Total reserve supply is equal to the quantity set by the central bank through its operations, adjusted by a potentially-random amount to reflect unpredictable changes in autonomous factors.

During the day, each bank makes payments to and receives payments from other banks. To keep things as simple as possible, suppose that each bank will make exactly one payment and receive exactly one payment during the "middle" part of the day. Furthermore, suppose that these two payment flows are of exactly the same size, $P_D > 0$, and that this size is non-stochastic. However, the order in which these payments occur during the day is random; some banks will receive the incoming payment before making the outgoing one, while others will make the outgoing payment before receiving the incoming one.

At the end of the day, after the interbank market has closed, each bank experiences another payment shock P that affects its end-of-day reserve balance. The value of P can be either positive, indicating a net outflow of funds, or negative, indicating a net inflow. We assume that the payment shock P is uniformly distributed on the interval $\left[-\overline{P}, \overline{P}\right]$. The value of this shock is not yet known when the interbank market is open; hence, a bank's demand for reserves in this market is affected by the distribution of the shock and not the realization.

We assume, as a starting point, that a bank must meet a given reserve requirement K at the end of each day.[5] If the bank finds itself holding fewer than K reserves at the end of the day,

[4] More specifically, we focus on the demand for nonborrowed reserve balances, that is, funds held by banks on deposit at the central bank that have not been borrowed from the central bank. For simplicity, in the remainder of the article we refer to these balances as "reserves."

[5] We discuss more complicated systems of reserve requirements later, including multiple-day maintenance periods. For the logic in the derivations that follow, the particular value of K does not matter. The case of $K = 0$ corresponds a system without reserve requirements.

after the payment shock P has been realized, it must borrow funds at a "penalty" rate of interest r_P to satisfy the requirement. This rate can be thought of as the rate charged by the central bank on discount window loans, adjusted to take into account any "stigma" associated with using this facility. In reality, a bank may pay a deficiency fee instead of borrowing from the discount window or it may borrow funds in the interbank market very late in the day when this market is illiquid. In the model, the rate r_P simply represents the cost associated with a late-day reserve deficiency, whatever the source of that cost may be.

The specific assumptions we make about the number and size of payments that a bank sends are not important; they only serve to keep the analysis free of unnecessary complications. Two basic features of the model are important. First, the bank cannot perfectly anticipate its end-of-day reserve position. This uncertainty creates a "precautionary" demand for reserves that responds to changes in the interest rate in a smooth way. Second, a bank makes payments during the day as a part of its normal operations and the pattern of these payments can potentially lead to an overdraft in the bank's reserve account. We initially assume that the central bank offers daylight credit to banks to cover such overdrafts at no charge. We study the case where daylight overdrafts are costly later in this section.

3.2 The benchmark case

We begin by analyzing a simple benchmark case; we show later in this section how the framework can be extended to include a variety of features that are important in reality. In the benchmark case, banks must meet their reserve requirement at the end of each day and the central bank pays no interest on reserves held by banks overnight. Furthermore, the central bank offers daylight credit free of charge.

Figure 1 depicts an individual bank's demand for reserves in this scenario. To draw this curve, we ask: Given a particular value for the interest rate, what quantity of reserves would the bank demand to hold if that rate prevailed in the market?

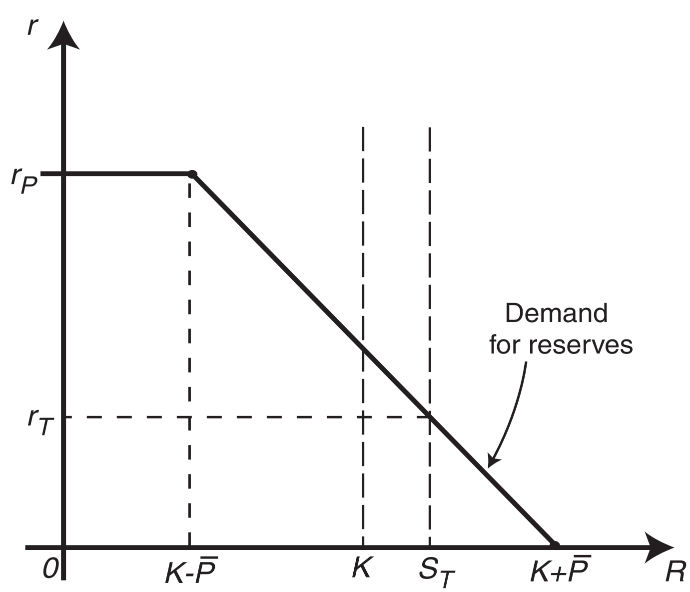

Figure 1: Benchmark demand for reserves

A bank would be unwilling to hold any reserves if the market interest rate were higher than r_P. If the market rate were higher than the penalty rate, the bank would choose to meet its requirement entirely by borrowing from the discount window. In fact, it would like to borrow even more than its requirement and lend the rest out at the higher market rate, but this fact is not important for the analysis. The important point is simply that there will be no demand for (nonborrowed) reserves for any interest rate larger than r_P.

When the market interest rate exactly equals the penalty rate r_P, a bank would be indifferent between holding any amount of reserves between zero and $K-\overline{P}$ and, hence, the demand curve is horizontal at r_P. As long as the bank's reserve holdings are lower than $K-\overline{P}$, the bank will need to borrow at the discount window to satisfy its reserve requirement K even if the late-day inflow of funds into the bank's reserve account is the largest possible value, $\overline{P}$. The alternative would be to borrow more reserves in the market to reduce this potential need for discount window lending. Since the market rate is equal to the penalty rate, both strategies deliver the same level of profit and the bank is indifferent between them.

For market interest rates below the penalty rate, however, a bank will choose to hold at least $K-\overline{P}$ reserves. As discussed above, if the bank held fewer than $K-\overline{P}$ reserves it would be certain to need to borrow from the discount window, which would not be an optimal choice when the market rate is lower than the discount rate. The bank's demand for reserves in this situation can be described as "precautionary" in the sense that the bank chooses its reserve holdings to balance the possibility of falling short of the requirement against the possibility of ending up with extra reserves in its account at the end of the day.

If the market interest rate were very low – close to zero – the opportunity cost of holding reserves would be very small. In this case, the bank would hold enough precautionary reserves so that it is virtually certain that unforeseen movements on its balance sheet will not decrease its reserves below the required level. In other words, the bank will hold $K+\overline{P}$ reserves in this case. If the market interest rate were exactly zero, there would be no opportunity cost of holding reserves. The demand curve is, therefore, flat along the horizontal axis after $K+\overline{P}$.

In between the two extremes, $K-\overline{P}$ and $K+\overline{P}$, the demand for reserves will vary inversely with the market interest rate measured on the vertical axis; this portion of the demand curve is represented by the downward-sloping line segment in Figure 1. The curve is downward-sloping for two reasons. First, the market interest rate represents the opportunity cost of holding reserves overnight. When this rate is lower, finding itself with excess balances is less costly for the bank and, hence, the bank is more willing to hold precautionary balances. Second, when the market rate is lower, the relative cost of having to access the discount window is larger, which also tends to increase the bank's precautionary demand for reserves.

The linearity of the downward-sloping part of the demand curve results from the assumption that the late-day payment shock is uniformly distributed. With other probability distributions, the demand curve will be nonlinear, but its basic shape will remain unchanged. In particular, the points where the demand curve intersects the penalty rate r_P and the horizontal axis will be the same for any distribution with support $\left[-\overline{P},\overline{P}\right]$.[6]

[6] The *support* of the probability distribution is the set of values of the payment shock that are assigned positive probability. An explicit formula for the demand curve in the uniform case is derived in Ennis and Weinberg (2007). If the shock instead had an unbounded distribution, such as the normal distribution used by Whitesell (2006) and others, the demand curve would asymptote to the penalty rate and the horizontal axis but never intersect them.

3.3 The equilibrium interest rate

Suppose, for the moment, that there is a single bank in the economy. Then the demand curve in Figure 1 also represents the total demand for reserves. Let S denote the total supply of reserves, as jointly determined by the central bank's operations and autonomous factors. Then the equilibrium interest rate is determined by the height of the demand curve at the point S. As shown in the diagram, there is a unique level of reserve supply S_T that will generate a given target interest rate r_T.

Now suppose there are many banks in the economy, but they are all identical in the sense that they have the same level of required reserves, face the same payment shock, etc. When there are many banks, the total demand for reserves can be found by simply "adding up" the individual demand curves. For any interest rate r, total demand is simply the sum of the quantity of reserves demanded by each individual bank.

For presentation purposes, it is useful to look at the average demand for reserves, that is, the total demand divided by the number of banks. When all banks are identical, the average demand is exactly equal to the demand of each individual bank. In other words, in the benchmark case where banks are identical, the demand curve in Figure 1 also represents the *aggregate* demand for reserves, expressed in per-bank terms. The determination of the equilibrium interest rate then proceeds exactly as in the single-bank case. In particular, the market-clearing interest rate will be equal to the target rate r_T if and only if reserve supply (expressed in per-bank terms) is equal to S_T.

Note that the central bank has two distinct ways in which it can potentially affect the market interest rate: changing the supply of reserves and changing (either directly or indirectly) the penalty rate. Suppose, for example, that the central bank wishes to decrease the market interest rate. It could either increase the supply of reserves through open market operations, leading to a movement down the demand curve, or it could decrease the penalty rate, which would rotate the demand curve downward while leaving the supply of reserves unchanged. Both policies would cause the market interest rate to fall.

3.4 Heterogeneity

While the assumption that all banks are identical was useful for simplifying the presentation above, it is clearly a poor representation of reality in most economies. The United States, for example, has thousands of banks and other depository institutions that differ dramatically in size, range of activities, etc. We now show how the analysis above changes when there is heterogeneity among banks and, in particular, how the size distribution of banks might affect the aggregate demand for reserves.

Each bank still has a demand curve of the form depicted in Figure 1, but now these curves can be different from each other because banks may different levels of required reserves, face different distributions of the payment shock, and/or face different penalty rates. Aggregating these individual demand curves can be done exactly as before: for any interest rate r, the total demand for reserves is simply the sum of the quantity of reserves demanded by each individual bank. The aggregate demand curve, expressed in per-bank terms, will again be similar to that presented in Figure 1, with the exact shape being determined by the properties of the various individual demands. If different banks have different levels of required reserves, for example, the requirement K in the aggregate demand curve will be equal to the average of the individual banks' requirements.

Our interest here is in studying how bank heterogeneity affects the properties of this demand curve. We focus on heterogeneity in bank size, which is particularly relevant in the U.S., where there are some very large banks and thousands of smaller banks. We ask how large banks may differ from small banks in the context of the simple framework and how the presence of both large and small banks might affect the properties of the aggregate demand curve. To simplify the presentation we study the three possible dimensions of heterogeneity addressed by the model one at a time. In reality, of course, the three cases will appear closely intertwined.

Size of Requirements. Perhaps the most natural way of capturing differences in bank size is by allowing for heterogeneity in reserve requirements. When requirements are calculated as a percentage of the deposit base, larger banks will tend to have a larger level of required reserves in absolute terms. Suppose, then, that banks have different levels of K, but they face the same late-day payment shock and the same penalty rate for a reserve deficiency. How would the size distribution of banks affect the aggregate demand for reserves in this case?

First note that, going back to Figure 1, the slope of the demand curve is independent of the size of the bank's reserve requirement K. To see why this is the case, consider an increase in the value of K. Since both $K - \overline{P}$ and $K + \overline{P}$ become larger numbers, the demand curve in Figure 1 shifts to the right. Notice that these two points shift exactly the same distance, leaving the slope of the downward-sloping segment of the demand curve unchanged.

Simple aggregation then shows that the slope of the aggregate demand curve will be independent of the size distribution of banks. In other words, for the case of heterogeneity in K, the sensitivity of reserve demand to changes in the interest rate does not depend at all on whether the economy is comprised of only large banks or, as in the U.S., has a few large banks and very many small ones.

Adding heterogeneity in reserve requirements does generate an interesting implication for the distribution of excess reserve holdings across banks. If large and small banks face similar (effective) penalty rates and are not too different in their exposure to late-day payment uncertainty, then the framework suggests that all banks should hold similar quantities of precautionary reserves. In other words, for a given level of the interest rate, the difference between the chosen reserve balances R and the requirement K should be similar for all banks. After the payment shocks are realized, of course, some banks will end up holding excess reserves and others will end up needing to borrow. On average, however, a large bank and a small one should end up holding comparable levels of excess reserves. If the banking system is composed of a relatively small number of large banks and a much larger number of small banks, then the majority of the excess reserves in the banking system will be held by small banks, simply because there are so many more of them. Even if large banks hold the majority of *total* reserve balances because of their larger requirements, most of the *excess* reserve balances will be held by small banks. This implication is broadly in line with the data for the U.S.

The Penalty Rate. Another way in which small banks might differ from large ones is the penalty rate they face if they need to borrow to avoid a reserve deficiency. To be eligible to borrow at the discount window, for example, a bank must establish an agreement with its Reserve Bank and post collateral. This fixed cost may lead some smaller banks to forgo accessing the discount window and instead borrow at a very high rate in the market (or pay the reserve deficiency fee) when necessary. Smaller banks may also have fewer established relationships with counterparties in the fed funds market and, as a consequence, may find it more difficult to borrow at a favorable interest rate late in the day (see Ashcraft, McAndrews and Skeie, 2007).

Suppose small banks do face a higher penalty rate, such as the value r_P^S depicted in Figure 2a, while larger banks face a lower rate r_P^L. The figure is drawn as if the two banks have the same level of requirements, but this is only to make the comparison between the curves clear. The figure shows two immediate implications of this type of heterogeneity. First, at any given interest rate, small banks will hold a higher level of precautionary reserves, that is, they will choose a larger reserve balance relative to their level of required reserves. In the figure, the smaller bank will hold a quantity S_S while the larger bank holds only S_L, even though – in this example – both face the same requirement and the same uncertainty about their end-of-day balance. As a result, the distribution of excess reserves in the economy will tend to be skewed even more heavily toward small banks than the earlier discussion would suggest.

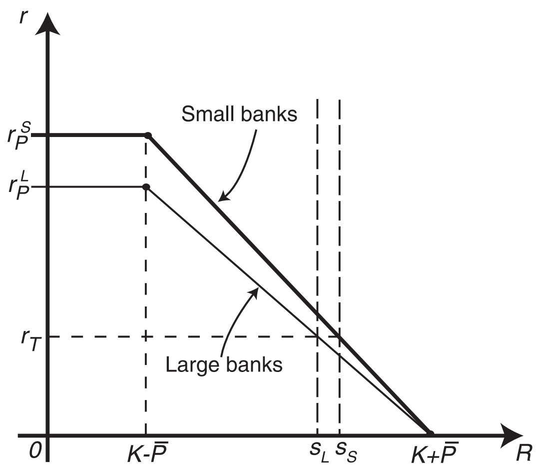

Figure 2a

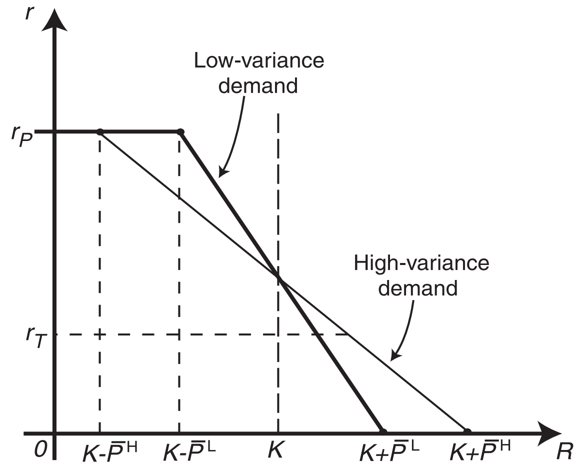

Figure 2b

The second implication shown in Figure 2a is that the demand curve for small banks has a steeper slope. In an economy with a large number of small banks, therefore, the aggregate demand curve will tend to be steeper, meaning that average reserve balances will be less sensitive to changes in the market interest rate. Notice that this result obtains even though there are no costs of reserve management in the model.

Support of the Payment Shock. A third way in which banks potentially differ from each other is the distribution of the late-day payment shock they face. Figure 2b depicts two demand curves, one for a bank facing a higher variance of this distribution and one for a bank facing a lower variance. The figure shows that having more uncertainty about the end-of-day reserve position leads to a flatter demand curve and, hence, a reserve balance that is more responsive to changes in the interest rate.

In this case, it is not completely clear which curve corresponds better to large banks and which to small banks. Banks with larger and more complex operations might be expected to face much larger day-to-day variations in their payment flows. However, such banks also tend to have sophisticated reserve management systems in place. As a result, it is not clear whether the end-of-day uncertainty faced by a large bank is higher or lower than that faced by a small bank.[7] The effect of the size distribution of banks on the shape of the aggregate demand curve is, therefore, ambiguous in this case.

[7] One possibility is that large banks face a wider support of the shock due to their larger operations, but face a smaller variance due to economies of scale in reserve management. This distinction cannot be captured in the figures here, which are drawn under the assumption that the distribution of the payment shock is uniform. For other distributions, the variance generally plays a more important role in the analysis than the support.

3.5 Daylight credit fees

So far, we have proceeded under the assumption that banks are free to hold negative balances in their reserve accounts during the day and that no fees are associated with such daylight overdrafts. Most central banks, however, place some restriction on banks' access to overdrafts. In many cases, banks must post collateral at the central bank in order to be allowed to overdraft their account. The Federal Reserve currently charges an explicit fee for daylight overdrafts to compensate for credit risk. We now investigate how reserve demand changes in the basic framework when access to daylight credit is costly.

Suppose a bank sends its daytime payment P_D before receiving the incoming payment. If P_D is larger than R (the bank's reserve holdings), the bank's account will be overdrawn until the offsetting payment arrives. Let r_e denote the interest rate the central bank charges on daylight credit, δ the time period between the two payment flows during the day, and π the probability that a bank sends the outgoing payment before receiving the incoming one. Then the bank's expected cost of daylight credit is $\pi r_e \delta \left(P_D - R\right)$. This expression shows that an additional dollar of reserve holdings will decrease the bank's expected cost of daylight credit by $\pi r_e \delta$. In this way, the terms at which the central bank offers daylight credit can influence the bank's choice of reserve position.[8]

Figure 3 depicts a bank's demand for reserves when daylight credit is costly (that is, when $r_e > 0$). The case studied in Figure 1 (that is, when $r_e = 0$) is included in the figure for reference. It is still true that there will be no demand for reserves if the market rate is above the penalty rate r_P. The interest rate measured on the vertical axis is (as in all of our figures) the rate for a 24-hour loan. If the market rate were above the penalty rate, a bank would prefer to lend out all of its reserves at the (high) market rate and satisfy its requirements by borrowing at the penalty rate. By arranging these loans to settle at approximately the same time on both days, this plan would have no effect on the bank's daylight credit usage and hence would generate a pure profit.

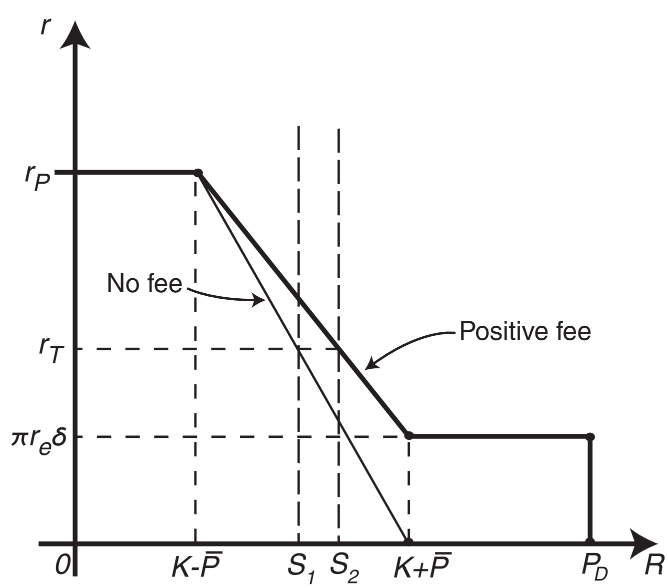

Figure 3: Daylight credit fees

[8] The treatment of overnight reserves can, in turn, influence the level of daylight credit usage. See Ennis and Weinberg (2007) for an investigation of this effect in a closely-related framework. See also the discussion in Keister, Martin, and McAndrews (2007).

It is also still true that whenever the market rate is below the penalty rate the bank will choose to hold at least $K - \overline{P}$ reserves, since otherwise it would be certain to need a discount window loan to meet its requirement. As the figure shows, the downward-sloping part of the demand curve is flatter when daylight credit is costly. For any market interest rate below the discount rate, the bank will choose to hold a higher quantity of reserves because these reserves now have the added benefit of reducing daylight credit fees.

Rather than decreasing all the way to the horizontal axis as in Figure 1, the demand curve now becomes flat at the bank's expected marginal cost of intraday funds, $\pi r_e \delta$. As long as R is smaller than P_D, the bank would not be willing to lend out funds at an interest rate below $\pi r_e \delta$ because the expected increase in daylight credit fees would be more than the interest earned on the loan. For values of R larger than P_D, the bank is holding sufficient reserves to cover all of its intraday payments and the demand curve drops to the horizontal axis.[9]

As the figure shows, when daylight credit is costly the level of reserves required to implement a given target rate is higher (S_2 rather than S_1 in the diagram). In other words, costly daylight credit tends to increase banks' reserve holdings. The demand curve is also flatter, meaning that reserve holdings are more sensitive to changes in the interest rate.

4 Interest rate volatility

One of the key determinants of a central bank's ability to consistently achieve its target interest rate is the slope of the aggregate demand curve for reserves. In this section, we describe the relationship between this slope and the volatility of the market interest rate in the basic framework. The next two sections then discuss policy tools that can be used to limit this volatility.

While the central bank can use open market operations to affect the supply of reserves available in the market, it typically cannot completely control this supply. Payments into and out of the Treasury account, as well as changes in the amount of cash in circulation, also affect the total supply of reserves. The central bank can anticipate much of the change in such autonomous factors, but there will often be significant unanticipated changes that cause the total supply of reserves to be different from what the central bank intended. As is clear from Figure 1, if the supply of reserves ends up being different from the intended amount S_T, the market interest rate will deviate from the target rate r_T.

Figure 4 illustrates the fact that a flatter demand curve for reserves is associated with less volatility in the market interest rate, given a particular level of uncertainty associated with autonomous factors. Suppose this uncertainty implies that, after a given open market operation, the total supply of reserves will be equal to either S or S' in the figure. With the steeper (thick) demand curve, this uncertainty about the supply of reserves leads to a relatively wide range of uncertainty about the market rate. With the flatter (thin) demand curve, in contrast, the variation in the market rate is smaller. For this reason, the slope of the demand curve, and those policies that affect the slope, are important determinants of the observed degree of volatility of the market interest rate around the target.

[9] The analysis here assumes a particular form of daylight credit usage; if an overdraft occurs, the size of the overdraft is constant over time. Alternative assumptions about the process of daytime payments would lead to minor changes in the figure, but the qualitative properties would be largely unaffected. The analysis also takes the size and timing of payments as given. Several papers have studied the interesting question of how banks respond to incentives in choosing the timing of their outgoing payments and, hence, their daylight credit usage. See, for example, McAndrews and Rajan (2000) and Bech and Garratt (2003).

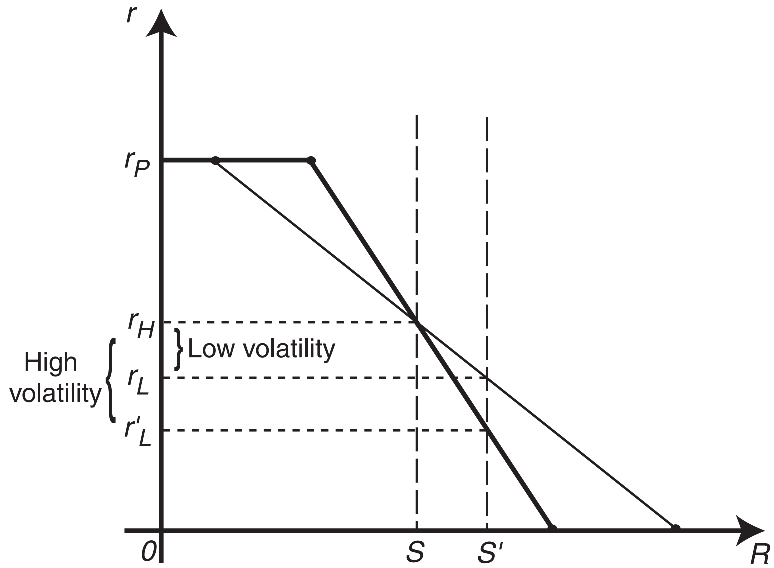

Figure 4: Interest rate volatility

As discussed in the previous section, a variety of factors affect the slope of the aggregate demand for reserves. Figure 4 can be viewed, for example, as comparing a situation where all banks face relatively little late-day uncertainty with one where all banks face more uncertainty; the latter case corresponds to the thin line in the figure. However, it should be clear that the reasoning presented above does not depend on this particular interpretation. The exact same results about interest rate volatility would obtain if the demand curves had different slopes because banks faced different penalty rates in the two scenarios or because of some other factor(s). What the figure shows is that there is a direct relationship between the slope of the demand curve and the amount of interest rate volatility caused by forecast errors or other unanticipated changes in the supply of reserves.

Central banks generally aim to limit the volatility of the interest rate around their target level to the extent possible. For this reason, a variety of policy arrangements have been designed in an attempt to decrease the slope of the demand curve, at least in the region that is considered "relevant". In the remainder of the article, we show how some of these tools can be analyzed in the context of our simple framework. In Section 4 we discuss reserve maintenance periods, while in Section 5 we discuss approaches that become feasible when the central bank pays interest on reserves.

5 Reserve maintenance periods

Perhaps the most significant arrangement designed to flatten the demand curve for reserves is the introduction of reserve maintenance periods. In a system with a reserve maintenance period, banks are not required to hold a particular quantity of reserves each day. Rather, each bank is required to hold a certain *average* level of reserves over the maintenance period. In the U.S., the length of the maintenance period is currently two weeks.

The presence of a reserve maintenance period gives banks some flexibility in determining when they hold reserves to meet their requirement. In general, banks will try to hold more reserves on days in which they expect the market interest rate to be lower and fewer reserves on days when they expect the rate to be higher. This flexibility implies that a bank's reserve holdings will tend to be more responsive to changes in the interest rate on any given day. In other words, having a reserve maintenance period tends to make the demand curve flatter, at least on days prior to the last day of the maintenance period. We illustrate this effect by studying a two-day maintenance period in the context of the simple framework. We then briefly explain how the same logic applies to longer periods.

5.1 A two-day maintenance period

Let K denote the average daily requirement, so that the total requirement for the two-day maintenance period is $2K$. The derivation of the demand curve for reserves on the second (and final) day of the maintenance period follows exactly the same logic as in our benchmark case. The only difference with Figure 1 is that the reserve requirement will be given by the amount of reserves that the bank has left to hold in order to satisfy the requirement for the period. In other words, the reserve requirement on the second day is equal to $2K$ minus the quantity of reserves the bank held at the end of the first day.

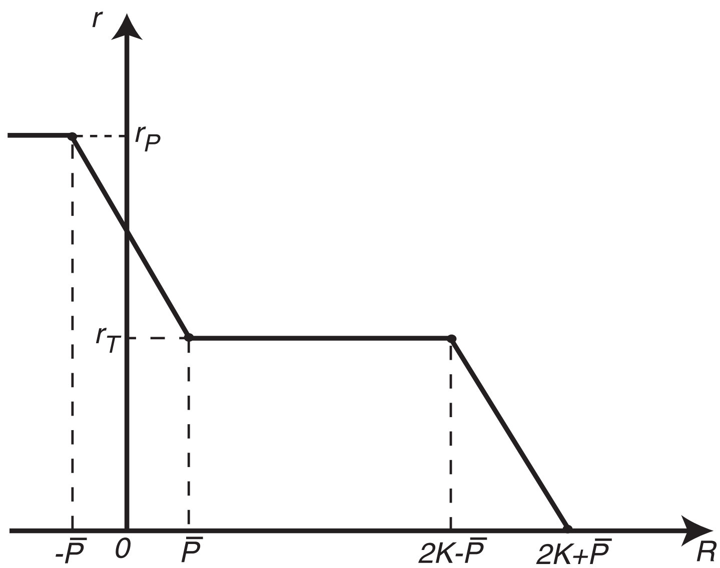

Figure 5: Reserve maintenance period

On the first day of the maintenance period, a bank's demand for reserves depends crucially on its belief about what the market interest rate will be on the second day. Suppose the bank expects the market interest rate on the second day to equal the target rate r_T. Figure 5 depicts the demand for reserves on the first day under this assumption.[10] As in the basic case presented in Figure 1, there would be no demand for reserves if the market interest rate were greater than r_P. Suppose instead that market interest rate on the first day is close to, but smaller than, the penalty rate r_P. Then the bank will want to satisfy as much of its reserve requirement as possible on the second day, when it expects the rate to be substantially lower. However, if the bank's reserve balance after the late-day payment shock is negative, it will be forced to borrow funds at the penalty rate to avoid incurring an overnight overdraft. As long as the market rate is below the penalty rate, therefore, the bank will choose a reserve position of at least $-\overline{P}$. Note that this reserve position represents

[10] For simplicity, Figure 5 is drawn with no discounting on the part of the bank. The effect of discounting is very small and inessential for understanding the basic logic described here.

the amount of reserves held by the bank before the late-day payment shock is realized. Even if this position is negative, as it would be the case when the market rate is close to r_P in Figure 5, it is still possible that the bank will receive a late-day inflow of reserves such that the bank does not need to borrow funds at the penalty rate to avoid an overnight overdraft. However, if the bank were to chose a position smaller than $-\overline{P}$, it would be certain to need to borrow at the penalty rate, which cannot be an optimal choice as long as the market rate is lower.

For interest rates below r_P, but still larger than the target rate, the bank will choose to hold some "precautionary" reserves to decrease the probability that it will need to borrow at the penalty rate. This precautionary motive generates the first downward-sloping part of the demand curve in the figure. As long as the day-one interest rate is above the target rate, however, the bank will not hold more than $\overline{P}$ in reserves on the first day. By holding $\overline{P}$, the bank is assured that it will have a positive reserve balance after the late-day payment shock. If the bank were holding more than $\overline{P}$ on the first day, it could lend those reserves out at the (relatively high) market rate and meet its requirement by borrowing reserves on the second day, when the interest rate is expected to be at the (lower) target rate, yielding a positive profit. Hence, the first downward-sloping part of the demand curve must end at $\overline{P}$.

Now suppose the first-day interest rate is exactly equal to the target rate r_T. In this case, the bank expects the rate to be the same on both days and is, therefore, indifferent between holding reserves on either day for the purpose of meeting reserve requirements. In choosing its first-day reserve position, the bank will consider the following issues. First, it will choose to hold at least enough reserves to ensure that it will not need to borrow at the penalty rate at the end of the first day. In other words, reserve holdings will be at least as large as the largest possible payment $\overline{P}$.

The bank is willing to hold more reserves than $\overline{P}$ for the purpose of satisfying some of its requirement. However, it wants to avoid the possibility of over-satisfying the requirement on the first day (that is, becoming "locked-in"), since it must hold a non-negative quantity of reserves on the second day to avoid an overnight overdraft. This implies that the bank will not be willing to hold more than the total requirement ($2K$) minus the largest possible payment inflow ($\overline{P}$) on the first day. The demand curve is flat between these two points (that is, $\overline{P}$ and $2K - \overline{P}$), indicating that the bank is indifferent between the various levels of reserves in this interval.

Finally, suppose the market interest rate on the first day is smaller than the target rate. Then the bank wants to satisfy most of the requirement the first day, since it expects the market rate to be *higher* on the second day. In this case, the bank will hold at least $2K - \overline{P}$ reserves on the first day. If it held any less than this amount, it would be certain to have some requirement remaining on the second day, which would not be an optimal choice given that the rate will be higher on the second day. As the interest rate moves farther below the target rate, the bank will hold more reserves for the usual precautionary reasons. In this case, the bank is balancing the possibility of being locked-in after the first day against the possibility of needing to meet some of its requirement on the more-expensive second day. The larger the difference between the rates on the two days is, the larger the quantity the bank will choose to hold on the first day. This trade-off generates the second downward-sloping part of the demand curve.

The intermediate flat portion of the demand curve in Figure 5 can help to reduce the volatility of the interest rate on days prior to the settlement day. As long as movements in autonomous factors are small enough such that the supply of reserves stays in this portion of the demand curve, interest rates fluctuations will be minimal. For a central bank that is interested in minimizing volatility around its target rate, this represents a substantial improvement over the situation depicted in Figure

1.[11]

There are, however, some issues that make implementing the target rate through reserve maintenance periods more difficult than a simple interpretation of Figure 5 might suggest. First, the position of the flat portion of the demand curve at the exact level of the target rate depends on the central bank's ability to hit the target rate (on average) on settlement day. If banks expected the settlement-day interest rate to be lower than the current target, for example, the flat portion of the first-day demand curve would also lie below the target. This issue is particularly problematic when market participants expect the central bank's target rate to change during the course of a reserve maintenance period. A second difficulty is that the flat portion of the demand curve disappears on the settlement day and the curve reverts to that in Figure 1.[12] This feature of the model indicates why market interest rates are likely to be more volatile on settlement days.

5.2 Multiple-day maintenance periods

Maintenance periods with three or more days can be easily analyzed in a similar way. Consider, for example, the case of a three-day maintenance period with an average daily requirement equal to K. As before, suppose that the central bank is expected to hit the target rate on the subsequent days of the maintenance period and consider the demand for reserves on the first day. This demand will be flat between the points $\overline{P}$ and $3K - \overline{P}$. That is, the demand curve will be similar to that plotted in Figure 4, but the flat portion will be wider.

To determine the shape of the demand for reserves in the second day we need to know how much reserves the bank held in the first day of the maintenance period. Suppose the bank held R_1 reserves with $R_1 < 3K$. Then on the second day of the maintenance period the demand for reserves would be flat between the points $\overline{P}$ and $3K - R_1 - \overline{P}$. Hence, we see that as the bank approaches the final day of the maintenance period the flat portion of its demand curve is likely to become smaller, potentially opening the door to increases in interest rate volatility. For the interested reader, Bartolini, Bertola, and Prati (2002) provide a more thorough analysis of the implications of multiple-day maintenance periods on the behavior of the overnight market interest rate using a model similar to, but more general than, ours.

6 Paying interest on reserves

We now introduce the possibility that the central bank pays interest on the reserve balances held overnight by banks in their accounts at the central bank. As discussed in Section 1, most central banks currently pay interest on reserves in some form, and Congress has authorized the Federal Reserve to begin doing so in October 2011. The ability to pay interest on reserves gives a central bank additional policy tools that can be used to help minimize the volatility of the market

[11] It should be noted that Figure 5 is drawn under the assumption that the reserve requirement is relatively large. Specifically, $K > \overline{P}$ is assumed to hold, so that the total reserve requirement for the period $(2K)$ is larger than the width of the support of the late-day payment shock $(2\overline{P})$. If this inequality were reserved, the flat portion of the demand curve would not exist. In general, reserve maintenance periods are most useful as a policy tool when the underlying reserve requirements are sufficiently large relative to the end-of-day balance uncertainty.

[12] In practice, central banks often use carryover provisions in an attempt to generate a small flat region in the demand curve on a settlement day. Another alternative would be to stagger the reserve maintenance periods for different groups of banks. This idea goes back to, at least, the 1960s (see, for example, the discussion in Sternlight, 1964). One common argument against staggering is that it could make the task of predicting reserve demand more difficult. Whether the benefits of reducing settlement day variability outweigh the potential costs of staggering is difficult to determine.

interest rate and to help steer this rate to the target level. This tool can be especially useful during periods of financial distress. For example, during the recent financial turmoil, the fed funds rate has experienced increased volatility during the day and has, in many cases, collapsed to values near zero late in the day. As we will see below, the ability to pay interest on reserves allows the central bank to effectively put a floor on the values of the interest rate that can be observed in the market. Such a floor reduces volatility and potentially increases the ability of the central bank to achieve its target rate.

In this section, we describe two approaches to monetary policy implementation that rely on paying interest on reserves: an interest rate corridor and a system with clearing bands. We explain the basic components of each approach and how it tends to flatten the demand curve for reserves.

6.1 Interest rate corridors

One simple policy a central bank could follow would be to pay a fixed interest rate r_D on all reserve balances that a bank holds in its account at the central bank.[13] This policy places a floor on the market interest rate: no bank would be willing to lend reserves at an interest rate lower than r_D, since they could instead earn r_D by simply holding the reserves on deposit at the central bank. Together, the penalty rate r_P and the deposit rate r_D form a "corridor" in which the market interest rate will remain.[14]

Figure 6 depicts the demand for reserves under a corridor system. As in the earlier figures, there is no demand for reserves if the market interest rate is higher than the penalty rate r_P. For values of the market interest rate below r_P, a bank will choose to hold at least $K - \overline{P}$ reserves for exactly the same reason as in Figure 1: if it held a lower level of reserves, it would be certain to need to borrow at the penalty rate r_P. Also as before, the demand for reserves is downward sloping in this region. The big change from Figure 1 is that the demand curve now becomes flat at the deposit rate. If the market rate were lower than the deposit rate, a bank's demand for reserves would be essentially infinite, as it would try to borrow at the market rate and earn a profit by simply holding the reserves overnight.

The figure shows that, regardless of the level of reserve supply S, the market interest rate will always stay in the corridor formed by the rates r_P and r_D. The width of the corridor $r_P - r_D$ is then a policy choice. Choosing a relatively narrow corridor will clearly limit the range and volatility of the market interest rate. Note that narrowing the corridor also implies that the downward slopping part of the demand curve becomes flatter (to see this, notice that the boundary points $K - \overline{P}$ and $K + \overline{P}$ do not depend on r_P or r_D). Hence, the size of interest rate movement associated with a given shock to an autonomous factor is smaller, even when the shock is small enough to keep the rate within the corridor.

An interesting case to consider is that where the lending and deposit rates are set the same distance on either side of the target rate (x basis points above and below the target, respectively). This system is called *a symmetric corridor*. A change in policy stance that involves increasing the

[13] In practice, reserve balances held to meet requirements are often compensated at a different rate than those that are excess to a bank's requirement. For the process of monetary policy implementation, the interest rate paid on *excess* reserves is what matters; this is the rate we denote r_D in our analysis.

[14] A central bank may prefer to use a *lending facility* that is distinct from its discount window to form the upper bound of the corridor. Banks may be reluctant to borrow from the discount window, which serves as a lender of last resort, because they fear that others would interpret this borrowing as a sign of poor financial health. The terms associated with the lending facility could be designed to minimize this type of stigma effect and, thus, create a more reliable upper bound on the market interest rate.

target rate, then, effectively amounts to changing the levels of the lending and deposit rates, which shifts the demand curve along with them. The supply of reserves needed to maintain a higher target rate, for example, may not be lower. In fact – perhaps surprisingly – in the simple model studied here the target level of the supply of reserves would not change at all when the policy rate changes.

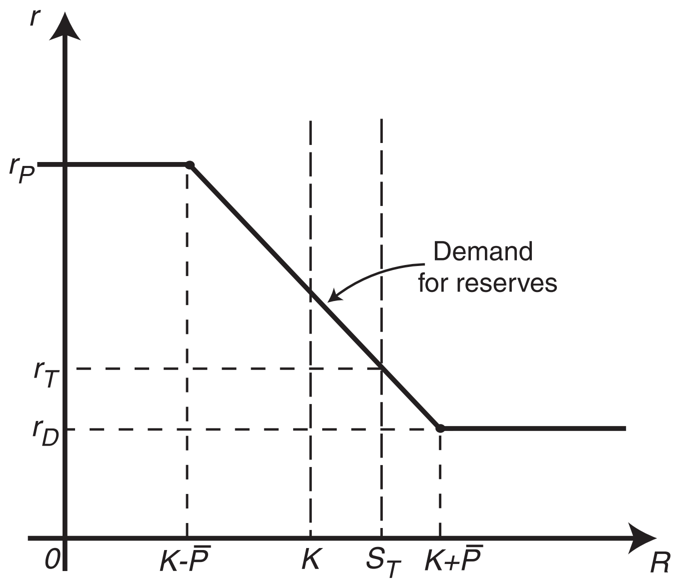

Figure 6: A conventional corridor

If the demand curve in Figure 6 is too steep to allow the central bank to effectively achieve its goal of keeping the market rate close to the target, a corridor system could be combined with a reserve maintenance period of the type described in Section 4. The presence of a reserve maintenance period would generate a flat region in the demand curve as in Figure 5. The features of the corridor would make the two downward-sloping parts of the demand curve in Figure 5 less steep, which would limit the interest rate volatility associated with events where reserve supply exits the flat region of the demand curve, as well as on the last day of the maintenance period when the flat region is not present.

Another way to limit interest rate volatility is for the central bank to set the deposit rate equal to the target rate and then provide enough reserves to make the supply S_T intersect the demand curve well into the flat portion of the demand curve at rate r_D. This "floor system" has been recently advocated as a way to simplify monetary policy implementation (see, for example, Woodford, 2000, Goodfriend, 2002 and Lacker, 2006). Note that such a system does not rely on a reserve maintenance period to generate the flat region of the demand curve, nor does it rely on reserve requirements to induce banks to hold reserves. To the extent that reserve requirements, and the associated reporting procedures, place significant administrative burdens on both banks and the central bank, setting the floor of the corridor at the target rate and simplifying, or even eliminating, reserve requirements could potentially be an attractive system for monetary policy implementation.

It should be noted, however, that the market interest rate will always remain some distance above the floor in such a system, since lenders in the market must be compensated for transactions costs and for assuming some counterparty credit risk. In other words, in a floor system the central bank is able to fully control the risk-free interest rate, but not necessarily the market rate. In normal times, the gap between the market rate and the rate paid on reserves would likely be stable and small.

In periods of financial distress, however, elevated credit risk premia may drive the average market interest rate significantly above the interest rate paid on reserves. Our simple model abstracts from these important considerations.[15]

6.2 Clearing bands

Another approach to generating a flat region in the demand curve for reserves is the use of daily clearing bands. This approach does not rely on a reserve maintenance period. Instead, the central bank pays interest on a bank's reserve holdings at the target rate r_T as long as those holdings fall within a pre-specified band. Let $\underline{K}$ and $\overline{K}$ denote the lower and upper bounds of this band, respectively. If the bank's reserve balance falls below $\underline{K}$, it must borrow at the penalty rate r_P to bring its balance up to at least $\underline{K}$. If, on the other hand, the bank's reserve balance is higher than $\overline{K}$, it will earn the target rate r_T on all balances up to $\overline{K}$ but will earn a lower rate r_E beyond that bound.

The demand curve for reserves under such a system is depicted in Figure 7. The figure is drawn under the assumption that the clearing band is fairly wide relative to the support of the late-day payment shock. In particular, we assume that $\underline{K} + \overline{P} < \overline{K} - \overline{P}$. Let us call the interval $\left[\underline{K} + \overline{P}, \overline{K} - \overline{P}\right]$ the "intermediate region" for reserves. By choosing any level of reserves in this intermediate region, a bank can ensure that its end-of-day reserve balance will fall within the clearing band. The bank would then be sure that it will earn the target rate of interest on all of the reserves it ends up holding overnight.

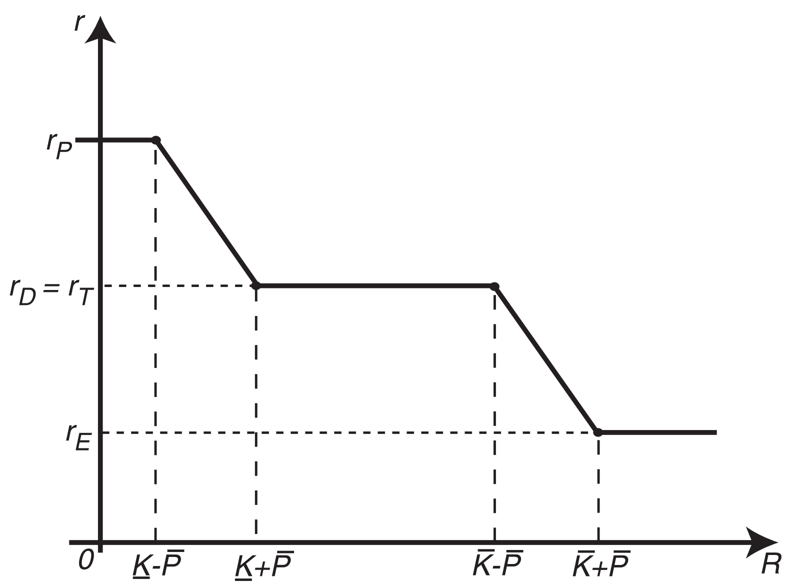

Figure 7: Required reserves clearing bands

When the market interest rate is equal to the target rate r_T, a bank is indifferent between choosing any level of reserves in the intermediate region. If the bank borrows in the market to slightly increase its reserve holdings, for example, the cost it would pay in the market for those reserves

[15] The central bank could also set an upper limit for the quantity of reserves on which it would pay the target rate of interest to a bank; reserves above this limit would earn a lower rate (possibly zero). Whitesell (2006) proposed that banks be allowed to choose their own upper limits by paying a facility fee per unit of capacity. Such an approach leads to a demand curve for reserves that is flat at the target rate over a wide region.

would be exactly offset by the extra interest it would earn from the central bank. Similarly, lending out reserves to slightly decrease the bank's holdings would also leave profit unchanged. This reasoning shows that the demand curve for reserves will be flat in the intermediate region between $\underline{K} + \overline{P}$ and $\overline{K} - \overline{P}$. As long as the central bank is able to keep the supply of reserves within this region, the market interest rate will equal the target rate r_T regardless of the exact level of reserve supply.

Outside the intermediate region, the logic behind the shape of the demand curve is very similar to that explained in our benchmark case. When the market interest rate is higher than r_T, a bank can earn more by lending reserves in the market than by holding them on deposit at the central bank. It would, therefore, prefer not to hold more than the minimum level of reserves needed to avoid being penalized, $\underline{K}$. Of course, the bank would be willing to hold some precautionary reserves to guard against the possibility that the late-day payment shock will drive their reserve balance below $\underline{K}$. The quantity of precautionary reserves it would choose to hold is, as before, an inverse function of the market interest rate; this reasoning generates the first downward-sloping part of the demand curve in Figure 7.

When the market rate is below r_T, on the other hand, the bank would like to take full advantage of its ability to earn the target interest rate by holding reserves at the central bank. It would, however, take into consideration the possibility that a late-day inflow of funds will leave it with a final balance higher than $\overline{K}$, in which case it would earn the lower interest rate r_E on the excess funds. The resulting decision process generates a downward-sloping region of the demand curve between the rates r_T and r_E. As in Figure 6, the demand curve never falls below the interest rate paid on excess reserves (now labeled r_E); this rate thus creates a floor for the market interest rate.

The demand curve in Figure 7 has the same basic shape as the one generated by a reserve maintenance period, which was depicted in Figure 4. It is important to keep in mind, however, that the forces generating the flat portion of the demand curve in the intermediate region are fundamentally different in the two cases. The reserve maintenance period approach relies on intertemporal arbitrage: banks will want to hold more reserves on days when the market interest rate is low and fewer reserves when the market rate is high. This activity will tend to equate the current market interest rate to the expected future rate (as long as the supply of reserves is in the intermediate region). The clearing band system relies instead on intraday arbitrage to generate the flat portion of the demand curve: banks will want to hold more reserves when the market interest rate is low, for example, simply to earn the higher interest rate paid by the central bank.

The intertemporal aspect of reserve maintenance periods has two clear drawbacks. First, if – for whatever reason – the expected future rate differs from the target rate r_T, it becomes difficult for the central bank to achieve the target rate in the current period. Second, large shocks to the supply of reserves on one day can have spillover effects on subsequent days in the maintenance period. If, for example, the supply of reserves is unusually high one day, banks will satisfy an unusually large portion of their reserve requirements and, as a result, the flat portion of the demand curve will be smaller on all subsequent days, increasing the potential for rate volatility on those days.

The clearing band approach, in contrast, generates a flat portion in the demand curve that always lies at the current target interest rate, even if market participants expect the target rate to change in the near future. Moreover, the width of the flat portion is "reset" every day; it does not depend on past events. These features are important potential advantages of the clearing band approach. We should again point out, however, that our simple model has abstracted from transaction costs and credit risk. As with the floor system discussed above, these considerations could result in the average market interest rate being higher than the rate r_T, as the latter represents a risk-free rate.

7 Conclusion

A recent change in legislation allowing the Federal Reserve to pay interest on reserves has renewed interest in the debate over the most effective way to implement monetary policy. In this article, we have provided a basic framework that can be useful for analyzing the main properties of the different alternatives. While we have conducted all our analysis graphically, our simplifying assumptions permit a fairly precise description of the alternatives and their effectiveness at implementing a target interest rate.

Many extensions of our basic framework are possible and we have analyzed several of them in this article. However, some important issues remain unexplored. For example, we only briefly mentioned the difficulties that fluctuations in aggregate credit risk can introduce in the implementation process. Also, as the debate continues, new questions will arise. We believe that the framework introduced in this article can be a useful first step in the search for the much needed answers to those questions.

Appendix A. Derivations

In this appendix, we provide a more formal derivation of the demand curves presented in the figures above. We formulate the profit function of a typical bank under each of the different policy regimes and derive the bank's optimal choice of reserve position. We also derive some properties of the resulting demand curve for reserves in each case.

A.1 The benchmark case

We begin with the benchmark case, which corresponds to Figure 1 in the text. Recall that, in this case, no interest is paid on reserve balances and there are no fees for daylight credit. If the bank's final reserve balance falls below the requirement $K,$ the difference must be borrowed at the penalty rate $r_P.$ Since r is the market interest rate, a bank's opportunity cost of holding a quantity R of reserve balances is given by the product $rR.$ The change in a typical bank's profits associated with its reserve operations can, therefore, be written as

$$\pi = -rR - \int_{R-K}^{\infty} r_P \left(P - (R - K)\right) f\left(P\right) dP,$$

where f is the density function for the late-day payment shock. Notice all of the terms in this expression are negative; when no interest is paid on reserve balances, reserve operations can only serve to lower a bank's profit. The bank is willing to incur these costs because it is required to hold reserves and make payments as a part of its (generally profitable) operations.

The bank will choose its reserve holdings R to maximize the value of π. The first-order condition for this problem is

$$\frac{\partial \pi}{\partial R} = -r + r_P \int_{R-K}^{\infty} f\left(P\right) dP = 0,$$

which can be solved for

$$r = r_P \left(1 - F\left(R - K\right)\right).$$

In other words, the optimal level of reserve balances equates the opportunity cost of holding one more unit of reserves with the marginal change in expected reserve deficiency costs. This latter change comes not from having a deficiency less often (which does happen, but is not a first-order effect), but rather from having a smaller deficiency when the payment shock P is high. The marginal change is, therefore, equal to the penalty rate r_P multiplied by the probability of a deficiency $\left(1 - F\left(R - K\right)\right).$ To put things slightly differently, the height of the demand curve in figure 1 is, for any given value of $R,$ equal to the marginal change in expected deficiency costs evaluated at $R.$

The slope of the demand curve in Figure 1 is given by

$$\frac{\partial r}{\partial R} = -r_P f\left(R - K\right). \tag{1}$$

This expression shows that the slope of the demand curve for reserves is proportional to the height of the density function for the payment shock. When the distribution of the shock is uniform, the slope of the demand curve is thus constant, as depicted in Figure 1. Under different distributional assumptions, the demand curve may have more "curvature", but the overall shape will remain similar. In particular, for any distribution with support $\left[-\overline{P}, \overline{P}\right]$, the demand curve will be flat at

r_P until the point $K - \overline{P}$ and will be flat on the horizontal axis after the point $K + \overline{P}$. Between these two points, the demand curve will always be downward sloping. Different distributions merely change the shape of this downward-sloping part of the curve.[16]

Suppose, for example, that the distribution of the late-day payment shock is hump-shaped, like the solid curve in the left panel of Figure 8. In this case, moderate values of P are more likely to occur than extreme values near either $-\overline{P}$ or $\overline{P}$. Using equation (1), it is easy to see that the corresponding demand curve must look like that depicted in the right panel of the figure, with a small slope for values near $K - \overline{P}$ and $K + \overline{P}$, but a steeper slope around the point K. Intuitively, because the probability of a payment shock near $-\overline{P}$ is very small, the bank is less concerned about a large payment inflow that would leave it holding excess reserves at the end of the day. As a result, the bank is willing to hold a larger quantity of reserves when the interest rate is high, which is why the demand curve in the right-hand panel lies above the dashed line for values of the overnight rate near r_P. The bank is also less concerned about a large payment outflow – that is, a realization near $\overline{P}$ – that might leave its end-of-day balance below K. It will choose, therefore, to hold fewer reserves than in the uniform case when the interest rate is near zero.

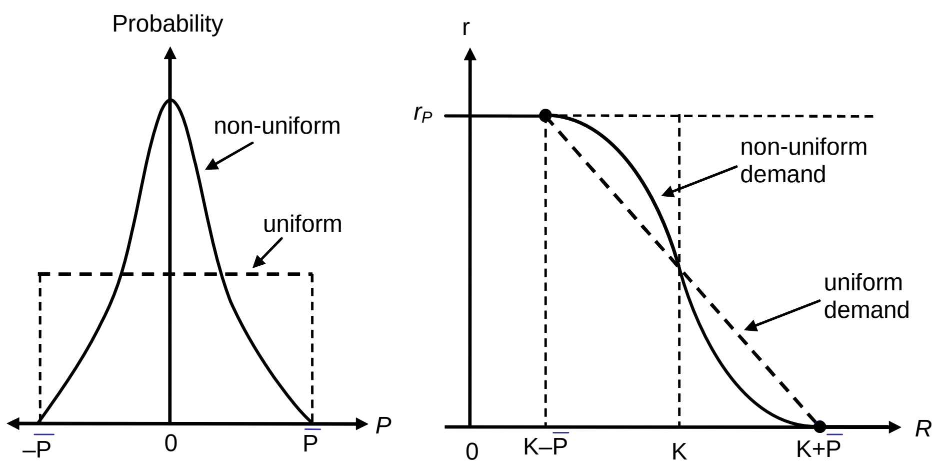

Figure 8: Slope of the demand curve

A.2 Interest rate corridors

Now suppose that the central bank remunerates reserve balances at a rate $r_D > 0$. In this case, a bank's profits associated with its reserve operations can be written as

$$\begin{aligned} \pi \;\; = \;\; & -rR + \int_{-\infty}^{R-K} r_D\left((R-K)-P\right) f(P)\, dP \\ & - \int_{R-K}^{\infty} r_P \left(P-(R-K)\right) f(P)\, dP + r_D K. \end{aligned} \tag{2}$$

The final term in this expression indicates that the bank must hold enough reserve balances to meet

[16] If the shock instead had an unbounded distribution, such as the normal distribution used by Whitesell (2006) and others, the demand curve would again have this same shape, but would asymptote to the rate r_P and to the horizontal axis without ever intersecting them.

its requirement K, and that it will earn interest at rate r_D on these balances. If, after the payment shock is realized, the bank is holding excess reserves, those will also be compensated at rate r_D; these situations are captured in the first integral in the equation. The second integral captures the situations where the shock is larger than $R - K$ and the bank must borrow at the penalty rate r_P to meet the requirement. Notice that reserves borrowed from the discount window and used to meet requirements are remunerated at the rate r_D and thus have a net cost of $(r_P - r_D)$.

The optimal reserve position of the bank is characterized by the first-order condition

$$\frac{\partial \pi}{\partial R} = -r + r_D \int_{-\infty}^{R-K} f(P)\, dP + r_P \int_{R-K}^{\infty} f(P)\, dP = 0.$$

The optimal choice now equates the opportunity cost of holding one more unit of reserves, r, with the marginal change in expected reserve deficiency costs *plus* the marginal change in expected interest income . Solving for the demand curve yields

$$r = r_D + (r_P - r_D)\,(1 - F(R - K))\,. \tag{3}$$

Here we see that the demand curve will never fall below the interest rate paid on reserves r_D, as depicted in Figure 6. The slope of the demand curve is given by

$$\frac{\partial r}{\partial R} = -\,(r_P - r_D)\, f(R - K)\,.$$

As in the benchmark case, we see that this slope is proportional to the height of the density function for the payments shock.

It is interesting to note that the interest rate paid on required reserves has no effect on the demand curve. This can be seen from the profit function (2), where the interest revenue from required reserves appears as a fixed, additively-separable payment. In the model studied here, where the reserve requirement is fixed independently of a bank's actions, remunerating reserves at a below-market rate simply acts as a lump-sum tax on banks and has no effect on bank behavior.

A.3 Reserve maintenance periods

We now examine the case of a two-day maintenance period, as studied in Section 5 above. We assume that excess reserves are remunerated at rate r_D and, for simplicity, that reserve balances held to meet requirements are not remunerated.[17] Let π_1 denote the net profit earned by the bank on the first day of the maintenance period, and let R_1 denote the bank's choice of reserve position on that day. Then we have

$$\pi_1 = -r_1 R_1 + \int_{-\infty}^{R_1 - 2K} r_D\,(R_1 - 2K - P)\, f(P)\, dP - \int_{R_1}^{\infty} r_P\,(P - R_1)\, f(P)\, dP, \tag{4}$$

where r_1 denotes the market interest rate on first day. If the bank experiences a large late-pay payment inflow $(P < R_1 - 2K)$, it will satisfy its entire requirement for the period on the first day. In this case, any reserves held beyond the required amount are remunerated at rate r_D. If the bank experiences a large late-day payment outflow $(P > R_1)$, the bank will be forced to borrow

[17] As discussed above, the remuneration rate on required reserves has no effect on the demand curves in our model.

at the penalty rate in order to avoid having an overnight overdraft. For intermediate values of the payments shock, however, the bank will neither have a deficit nor accumulate any excess reserves; its reserve balance at the end of the day is simply applied toward the total requirement.

Let R_2 denote the bank's reserve holdings on the second (and final) day of the maintenance period and r_2 the market interest rate on that day. Let π denote the total expected profit at the end of the maintenance period. Then we can write

$$\begin{aligned} \pi &= \pi_1 + r_2 (\pi_1 - R_2) + \int_{-\infty}^{R_2 - K_2} r_D (R_2 - K_2 - P) f(P) \, dP \\ &\quad - \int_{R_2 - K_2}^{\infty} r_P (P - (R_2 - K_2)) f(P) \, dP, \end{aligned} \tag{5}$$

where

$$K_2 = \left\{ \begin{array}{c} 2K \\ 2K - (R_1 - P_1) \\ 0 \end{array} \right\} \text{ if } R_1 - P_1 \left\{ \begin{array}{c} \leq 0 \\ \in (0, 2K) \\ \geq 2K \end{array} \right\}$$

and P_1 denotes the realization of the bank's payment shock on the first day. The variable K_2 measures the remaining requirement to be met on that day (if any), which typically equals the total requirement $2K$ minus the bank's end-of-day balance on the first day $(R_1 - P_1)$. Following the steps in the previous subsection, the demand curve on the last day of the maintenance period is easily seen to be

$$r_2 = r_D + (r_P - r_D)(1 - F(R_2 - K_2)).$$

Notice that is expression depends on first-day variables (r_1 and R_1) only through their effect on K_2. Also note that the bank's optimal choice of R_2 will move one-for-one with the remaining requirement K_2, that is, $dR_2/dK_2 = 1$ holds in the relevant region.

On the first day of the maintenance period, the bank will choose R_1 in order to maximize expected profits, given its belief about the interest rate that will prevail on the second day. Assume, for simplicity, that the bank has perfect foresight about the rate r_2. We have already shown that the choice of R_1 does not affect the difference $(R_2 - K_2)$. Therefore, this choice has no effect on the last two terms in the expression for total profit (5). In effect, then, the bank's reserve position on the first day is chosen to solve

$$\max_{R_1} \; \pi_1 + r_2 (\pi_1 - E[R_2(R_1; P_1)]),$$

where R_2 will be chosen optimally given K_2, which depends on R_1 and the realization of P_1. In other words, the bank chooses the quantity of reserves it holds on the first day to maximize its profit on the first day, taking into account the effect this choice will have on its reserve holdings on the second day. Using the solution for the second day derived above, we can show the relationship between R_1 and R_2 to be characterized by

$$\frac{dR_2}{dR_1} = \frac{dR_2}{dK_2} \frac{dK_2}{dR_1} = \left\{ \begin{array}{c} 0 \\ -1 \\ 0 \end{array} \right\} \text{ for } R_1 \left\{ \begin{array}{c} < P_1 \\ \in (P_1, P_1 + 2K) \\ > P_1 + 2K \end{array} \right\}.$$

Using this relationship and substituting in for π_1 from (4) yields

$$\max_{R_1} \ (1+r_2)\left(-r_1R_1+\int_{-\infty}^{R_1-2K} r_D\left(R_1-2K-P\right)f\left(P\right)dP-\int_{R_1}^{\infty} r_P\left(P-R_1\right)f\left(P\right)dP\right)$$
$$-r_2\left(\int_{-\infty}^{R_1-2K}\underline{R}_2\left(P\right)dP+\int_{R_1-2K}^{R_1} R_2\left(R_1;P\right)f\left(P\right)dP+\int_{R_1}^{\infty}\overline{R}_2 f\left(P\right)dP\right),$$

where $\underline{R}_2$ is the quantity of reserves the bank will choose to hold on the second day if $K_2=0$ and $\overline{R}_2$ is the corresponding quantity for $K_2=2K$. Both of these numbers are constants, independent of the choice of R_1.

The first-order condition for this problem can be written as

$$-r_1+r_D\int_{-\infty}^{R_1-2K} f\left(P\right)dP-r_P\int_{R_1}^{\infty} f\left(P\right)dP-\frac{r_2}{1+r_2}\int_{R_1-2K}^{R_1} f\left(P\right)dP\frac{dR_2}{dR_1}=0.$$

The first part of this expression is similar to the earlier first-order conditions: it reflects the opportunity cost of holding reserves, r_1, as well as the marginal changes in expected deficiency costs and expected interest earnings on the first day. The last term in the expression is new; it reflects the expected effect of first-day reserve holdings on second-day reserve holdings. This condition can be solved for the demand function

$$r_1=r_P-\left(r_P-\frac{r_2}{1+r_2}\right)F\left(R_1\right)-\left(\frac{r_2}{1+r_2}-r_D\right)F\left(R_1-2K\right). \qquad (6)$$

This function corresponds to the demand curve depicted in Figure 5.

To see why the demand curve in (6) generates the shape presented in Figure 5, first consider very low (i.e., negative) values of R_1. If R_1 is small enough, both $F\left(R_1\right)$ and $F\left(R_1-2K\right)$ will be zero (or very close to zero). From (6), the corresponding market interest rate would then be r_P. In other words, the demand curve is initially flat at the level r_P, as depicted in the figure. Next consider the other extreme case, where R_1 is large enough that both $F\left(R_1\right)$ and $F\left(R_1-2K\right)$ are close to unity. In this case, the corresponding market interest rate is equal to r_D; hence, the demand curve is eventually flat at level r_D, again as depicted in the figure. Finally, suppose that K is large enough so that for some intermediate values of R_1, we have both

$$F\left(R_1\right)\approx 1 \quad \text{and} \quad F\left(R_1-2K\right)\approx 0.$$

For these values of $R_{1,}$ the demand curve lies at

$$\frac{r_2}{1+r_2}\approx r_2.$$

Note the approximation here. In deriving Figure 5, we said that a bank would be indifferent between holding reserves on the two days if $r_1=r_2$ holds. This is not quite correct, since the bank should discount the opportunity cost of holding reserves on the second day. However, for reasonable values of the daily interest rate, this discounting is immaterial. (Formally, r_2 is the best first-order approximation of $r_2/\left(1+r_2\right)$ around the point $r_2=0$.) Hence, for intermediate values of R_1, the demand curve will be flat at a value very close to r_2 as long as the total requirement $2K$

is large enough. In such cases, the demand curve in (6) looks precisely like the one depicted in Figure 5.

A.4 Clearing Bands

Now suppose that a bank has a single-day reserve requirement with a clearing band of the type discussed in Section 6.2. The bank must hold a minimum reserve balance $\underline{K}$ at the end of the day, borrowing at the penalty rate if necessary to make up any deficiency. The bank will earn the target rate of interest r_T on all balances up to some limit $\overline{K} > \underline{K}$. Above $\overline{K}$, all reserves are remunerated at a lower rate r_D, which could be zero. In other words, the bank will earn the target rate of interest r_T on all of its reserves as long at the total falls in the clearing band $\left[\underline{K}, \overline{K}\right]$. Outside of this clearing band, the costs and benefits are set as in a channel system.[18]

A bank's expected profit associated with its reserve operations under this system is

$$\begin{aligned}\pi &= -rR + \int_{-\infty}^{R-\overline{K}} \left(r_T\overline{K} + r_D\left(R - P - \overline{K}\right)\right) f\left(P\right) dP \\ &\quad + \int_{R-\overline{K}}^{R-\underline{K}} r_T\left(R - P\right) f\left(P\right) dP + \int_{R-\underline{K}}^{\infty} \left(r_T\underline{K} - r_P\left(P - \left(R - \underline{K}\right)\right)\right) f\left(P\right) dP.\end{aligned}$$

The first integral in this expression captures situations where the late-pay payment shock is small enough that the bank's final reserve balance is greater than $\overline{K}$ (this might, for example, happen if the bank experiences a large late-day payment inflow). In such instances, the bank earns the rate r_T and the first $\overline{K}$ reserves and the rate r_D on the remainder. The second integral captures intermediate values of the payment shock, which leave the bank's final reserve balance between $\overline{K}$ and $\underline{K}$, in which case the bank earns the rate r_T on all of these balances. The third integral captures large payment outflows that leave the bank's final reserve balance below $\underline{K}$. In these cases, the bank must borrow at the penalty rate r_P to meet the minimal requirement $\underline{K}$.

As before, the bank will choose R in order to maximize expected profit. The first-order condition for this problem can be written as

$$\frac{\partial\pi}{\partial R} = -r + r_D \int_{-\infty}^{R-\overline{K}} f\left(P\right) dP + r_T \int_{R-\overline{K}}^{R-\underline{K}} f\left(P\right) dP + r_P \int_{R-\underline{K}}^{\infty} f\left(P\right) dP = 0.$$

Once again, the optimal choice of reserve position involves balancing the opportunity cost of holding reserves, r, against the marginal changes in both expected deficiency costs and expected interest receipts. Solving for the demand curve yields

$$r = r_D + \left(r_T - r_D\right)\left(1 - F\left(R - \overline{K}\right)\right) + \left(r_P - r_T\right)\left(1 - F\left(R - \underline{K}\right)\right). \tag{7}$$

This demand curve corresponds to the one presented in Figure 7. Its slope is given by

$$\frac{\partial r}{\partial R} = -\left(r_T - r_D\right) f\left(R - \overline{K}\right) - \left(r_P - r_T\right) f\left(R - \underline{K}\right).$$

The understand the shape of this curve, first consider values of R that are low enough that

[18] Note that if $\underline{K} = \overline{K}$, this system becomes a channel system with rate r_T paid on reserves held to meet requirements.

both $F\left(R-\overline{K}\right)$ and $F\left(R-\underline{K}\right)$ are zero (or very close to zero). In such cases, the interest rate emerging from (7) is the penalty rate r_P. In other words, the demand curve is initially flat at r_P. Next, consider very large values of R, so that both $F\left(R-\overline{K}\right)$ and $F\left(R-\underline{K}\right)$ are equal to unity. In these cases, the interest rate from (7) is r_D, meaning that the demand curve is eventually flat at this level. Finally, consider intermediate values of R. If the clearing band $\left[\underline{K},\overline{K}\right]$ is wide enough, there will exist some values of R such that

$$F\left(R-\overline{K}\right)\approx 0 \quad \text{and} \quad F\left(R-\underline{K}\right)\approx 1.$$

For these values, (7) shows that the demand curve will be flat at the target rate r_T, as depicted in Figure 7.

References

[1] Ashcraft, Adam, James J. McAndrews, and David Skeie (2007). "Precautionary Reserves and the Interbank Market," *mimeo* (July), Federal Reserve Bank of New York.

[2] Bartolini, Leonardo, Giuseppe Bertola, and Alessandro Prati (2002). "Day-To-Day Monetary Policy and the Volatility of the Federal Funds Interest Rate," *Journal of Money, Credit, and Banking* 34 (1, February), pp. 137-159.

[3] Bartolini, Leonardo, Svenja Gudell, Spence Hilton, and Krista Schwarz (2005). "Intraday Trading in the Overnight Federal Funds Market," Federal Reserve Bank of New York *Current Issues in Economics and Finance* 11 (11, November), pp. 1-7.

[4] Bech, M. L. and Rod Garratt (2003). "The Intraday Liquidity Management Game," *Journal of Economic Theory* 109 (2, April), pp. 198-219.

[5] Clouse, James A. and James P. Dow Jr. (2002). "A Computational Model of Banks' Optimal Reserve Management Policy," *Journal of Economic Dynamics & Control* 26 (11, September), pp. 1787-1814.

[6] Dotsey, Michael (1991). "Monetary Policy and Operating Procedures in New Zealand," Federal Reserve Bank of Richmond *Economic Review* (September/October), pp. 13-19.

[7] Ennis, Huberto M. and John A. Weinberg (2007). "Interest on Reserves and Daylight Credit," Federal Reserve Bank of Richmond *Economic Quarterly* 93 (2, Spring), pp. 111-142.

[8] Goodfriend, Marvin (2002). "Interest on Reserves and Monetary Policy," Federal Reserve Bank of New York *Economic Policy Review* 8 (1, May), pp. 1-8.

[9] Guthrie, Graeme and Julian Wright (2000). "Open Mouth Operations," *Journal of Monetary Economics* 46, pp. 489-516.

[10] Hilton, Spence and Warren B. Hrung (2007). "Reserve Levels and Intraday Federal Funds Rate Behavior," Federal Reserve Bank of New York Staff Report no. 284 (May).

[11] Lacker, Jeffrey M. (2006). "Central Bank Credit in the Theory of Money and Payments." *Speech*. http://www.richmondfed.org/

[12] Keister, Todd, Antoine Martin, and James J. McAndrews (2007). "Bank Reserves and Monetary Policy Implementation," *mimeo.*, Federal Reserve Bank of New York.

[13] Martin, Antoine and James J. McAndrews (2006). "Liquidity Saving Mechanisms," *mimeo.* (September), Federal Reserve Bank of New York.

[14] McAndrews, James J. (2006). "Alternative Arrangements for the Distribution of Intraday Liquidity," Federal Reserve Bank of New York *Current Issues in Economics and Finance* 12 (3, April), pp. 1-11.

[15] McAndrews, James J. and Samira Rajan (2000). "The Timing and Funding of Fedwire Funds Transfers," Federal Reserve Bank of New York *Economic Policy Review* 6, (2, July), pp. 17-32.

[16] Poole, William (1968). “Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy,” *Journal of Finance* 23 (December), pp. 769-791.

[17] Sternlight, Peter D. (1964) “Reserve Settlement Periods of Member Banks: Comment” *Journal of Finance* 19 (1, March), pp. 94-98.

[18] Whitesell, William (2006). “Interest Rate Corridors and Reserves,” *Journal of Monetary Economics* 53, pp. 1177-1195.

[19] Woodford, Michael (2000). “Monetary Policy in a World Without Money,” *International Finance* 3 (2, July), pp. 229-260.