Donald S. Allen is an economist at the Federal Reserve Bank of St. Louis. Thomas A. Pollmann provided research assistance.

How Closely Do Banks Manage Vault Cash?

Donald S. Allen

Beginning in 1959, Federal Reserve member banks were allowed to use both vault cash and deposits at Federal Reserve Banks to satisfy statutory reserve requirements. Prior to this time, member banks were restricted to satisfying required reserves solely with deposits at Federal Reserve Banks. Hence, member banks held currency in their vaults only to satisfy daily demands by depositors. Fluctuations in the levels of vault cash prior to 1959 therefore can be assumed to reflect changes in depositors' demands for cash, or changes in the opportunity cost of vault cash (the cost of holding versus the cost of running out). Figure 1 shows the historical movement of vault cash as a percent of demand deposits at depository institutions from 1930 to 1997. During the 1930s, vault cash as a percent of demand deposits ranged between 4 and 5 percent, possibly reflecting higher public demand for currency or more conservative management by banks as a result of the banking crises of the early 1930s and the Great Depression. During the period from the mid-1940s to the late 1950s, vault cash was typically around 2.5 percent of demand deposits.

After 1959, decisions on how much currency to hold in vaults became influenced by total reserve management decisions. Banks could choose the amount of their required reserves that they wanted to be satisfied by currency in the vaults and the amount they wanted to hold in their accounts at the Fed. The opportunity costs were reduced for holding vault cash in excess of that required for customer demand but less than required for reserves. Banks could hold more currency than they needed for customer transactions if this currency also satisfied a portion of required reserves. The result, as Figure 1 shows, was that vault cash as a percent of total demand deposits increased steadily from 1959.[1]

Depository institutions are currently in an era of relatively low reserve requirements, yet bank vault cash holdings have been increasing—both in magnitude and as a percent of net transaction deposits.

Varying Requirements for Nonmember Banks

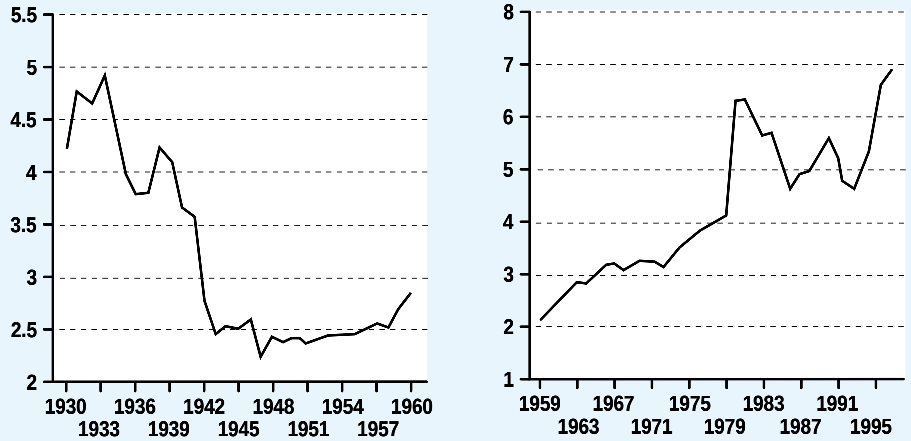
The role of vault cash for nonmember banks and thrifts is more complex than that for Federal Reserve member banks, as described in this article. Prior to the Monetary Control Act of 1980, these institutions were subject to a wide range of state-imposed reserve requirements (see Gilbert and Lovati, 1978). In some states, no requirements were imposed and vault cash presumably was held solely to meet customer needs. In others, vault cash could be applied to satisfy required reserves. This system changed dramatically in November 1980, when all nonmember banks and thrifts that offer transaction deposits became subject to the same reserve requirements as member banks and could use vault cash to satisfy such requirements. The requirements for nonmember institutions were phased in over a seven-year period from 1980 to 1987. During this period, the vault cash behavior of nonmember institutions likely was governed by a complicated mixture of customer needs and statutory reserve requirements.

[1] These data are for all depository institutions. Separating Federal Reserve member banks introduces breaks in the data. We can assume that the changes in member bank reserve management methods contributed to the increase. Other checkable deposits that became significant during the 1980s are included as part of demand deposits for this period.

Figure 1

Vault Cash as a Percent of Demand Deposits

All Depository Institutions, End of Year



SOURCE: Friedman and Schwartz–A Monetary History of the United States, 1867-1960

SOURCE: Federal Reserve Data (Demand Deposits plus Other Checkable Deposits)

In December 1990 and April 1992, reserve requirements were reduced significantly. In addition, starting in 1995, banks began "sweeping" customers' cash reserves into savings accounts that were not subject to reserve requirements, further reducing the amounts designated as required reserves (and reported net transaction deposits). Because of this reduction, the percentage of required reserves satisfied by vault cash has increased steadily. At the same time, the number of "bound banks," or banks whose required reserves are not satisfied completely by vault cash, has decreased.

As shown in Table 1, the number of bound small banks in the Eighth Federal Reserve District decreased from 27 percent in 1993 to 20 percent in 1996; the number of bound medium banks decreased from 74 percent in 1993 to 51 percent in 1996; and the number of bound large banks decreased from 100 percent in 1993 to 72 percent in 1996. In addition, the amount of total reserve requirements satisfied entirely by vault cash increased from 54 percent in 1993 to 75 percent in 1996.

Figure 2 shows total required reserves and the total vault cash held by banks in the Eighth District. The downward trend in reserves over the last three years, accompanied by an upward trend in vault cash, is obvious. By itself, the increase in vault cash does not imply a surplus. But, as Figure 3 (taken from Anderson and Rasche, 1996) shows, the quantity of "surplus" vault cash (i.e., above that used to meet reserve requirements) has increased steadily. This rise in surplus vault cash, in conjunction with the increase in nonbound banks, suggests that vault cash levels are not managed very closely or are constrained by precautionary balances to serve customer needs, rather than statutory reserve requirements.

This article looks at daily vault cash holdings in the Eighth District (headquartered in St. Louis) to determine whether the observed amounts of vault cash held by banks are consistent with the fundamental assumptions of a one-sided (*S*,*s*) inventory decision rule, a well-known model of optimal firm inventory management.[2] It also reviews how the limits implied by the (*S*,*s*) model change as a function of the mean and variance of demand and the opportunity cost of maintaining inventories.

[2] An (*S*,*s*) inventory decision rule establishes an upper limit, *S*, and a lower limit, *s*, on inventory. When stocks fall below *s*, inventories are replenished to the upper limit.

A Closer Look at the Reserve Requirements Picture

As of 1992, the required reserve margins established by the Federal Reserve for depository institutions were 3 percent of net transaction deposits up to $49.3 million and 10 percent on net transaction deposits above $49.3 million. The rate on the second tranche was reduced from 12 percent to 10 percent in April 1992, and the size of the first tranche was reduced from $52.0 million to $49.3 million in January 1997. Larger depository institutions have either a weekly or biweekly reserve account maintenance period, beginning on a Thursday and ending on a Wednesday, and reserve requirements are on a contemporaneous basis. The typical reserve requirement therefore depends on the size of the bank. The reserve requirement in the Eighth District averages to about 8 percent of transaction deposits, with smaller banks averaging around 3 percent and larger banks averaging more than 8 percent. Reserves are held as vault cash or deposits at the Federal Reserve Bank. Deposits at the Federal Reserve that are not used to satisfy reserve requirements may be used to satisfy clearing balances and are available for overnight loans to other institutions.

The rest of the paper is organized as follows: The first section following the overview looks at why banks hold vault cash and what parameters affect how much they hold. The next section reviews the vault cash data for the Eighth District to determine how recent developments in reserve requirements have affected vault cash levels. The following section provides a review of the theoretical motivation for (*S,s*) inventory behavior, and simulations are performed to provide insight into whether banks are managing vault cash

Table 1

Percent of Banks with Required Reserves Satisfied by Vault Cash

Bank Size	1993	1994	1995	1996
Small	73	73	79	80
Medium	26	33	43	49
Large	0	5	19	28
Total Reserve Requirement Satisfied by Vault Cash	54%	57%	62%	75%

Figure 2

Total Required Reserves vs. Total Vault Cash

Eighth Federal Reserve District (Bi-Weekly, January 1989 - December 1996)

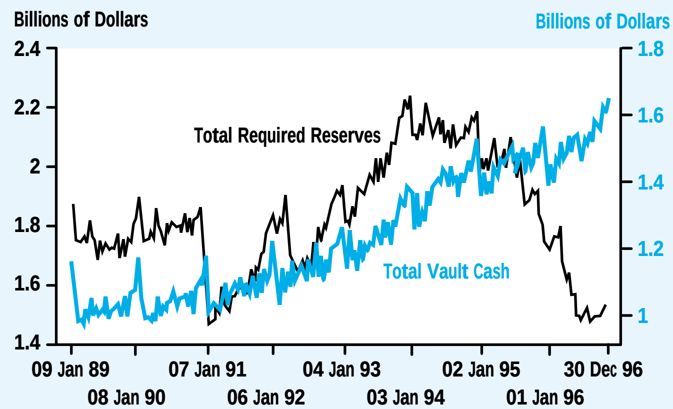
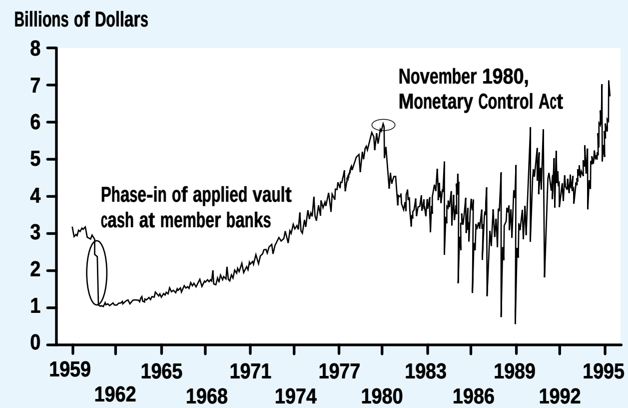


Figure 3

Surplus Vault Cash



conservatively. The final section offers some tentative conclusions.

WHY DO BANKS HOLD VAULT CASH?

Banks maintain an inventory of currency in vaults primarily to meet the daily currency demands of depositors, either at the teller window or in automated teller machines (ATMs). If banks knew with certainty the daily flows of currency from deposits and withdrawals, the amount of vault cash required could be determined precisely. In addition, if there were no costs associated with the shipment of currency, banks could reduce their holdings of vault currency and increase the frequency of currency shipments. However, the relative randomness of currency flows and the fixed cost of obtaining currency shipments induce banks to hold higher levels of vault cash. A bank's incentive to minimize vault cash holdings is reduced further when such cash also counts as part of required reserves.

Uncertainty in the daily flow of currency dictates that banks hold some average level of vault currency to cover the probability of "stocking out" of cash. The fixed costs associated with currency shipments make it cost effective to achieve this average vault cash level by infrequent shipments of larger amounts of currency. Thus, banks identify some trigger level below which vault cash will be replenished and either a maximum level to which vault cash will be replenished or an economic batch amount, which is considered an optimum shipment amount. This situation leads to a decision rule referred to in the inventory management literature as (S,s) behavior, where S is the upper limit of vault cash, and s is the lower limit or replenishment signal. Both the upper and lower limits will depend on the intraday profile of withdrawals and deposits, as well as the cost associated with shipments and the opportunity costs of stocking out.

The intraday demand profile is a function of the types of customers the bank serves. For example, a bank that primarily serves retail firms may see a net daily inflow of currency, whereas a bank that primarily serves individual customers may receive the bulk of its deposits in the form of checks and thus experience a net outflow of currency. If the bank experiences a net inflow of currency, which may be the case if its customers are primarily commercial enterprises, the bank may identify an upper limit of vault cash that will induce it to send currency back to the Fed to credit to its account. If a bank experiences a net currency withdrawal each day, then a large shipment may be received each time the level falls below the trigger, boosting the level of vault cash. This level is drawn down over a period of time (days) until the lower level is reached and another shipment is received, again boosting vault cash levels. Both types will lead to a sawtooth pattern of vault cash levels. In one case, vault cash will rise slowly and fall quickly; in the other, vault cash will rise quickly and fall gradually.

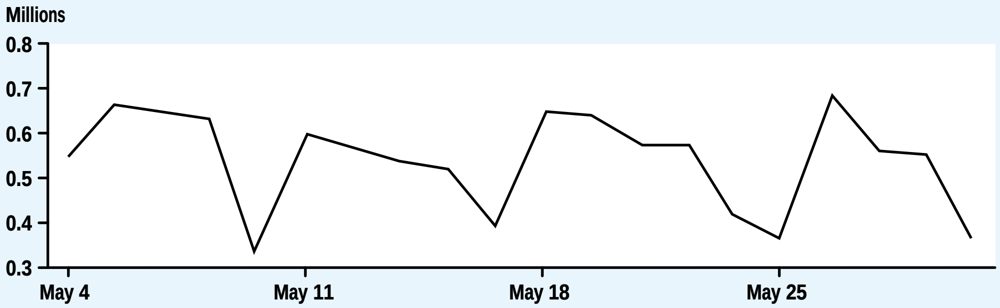
Figures 4, 5, and 6 show this typical pattern for a small, medium, and large bank in the Eighth District over a period in 1992. The small bank's vault cash rises on Mondays, possibly from shipments of currency from the Fed or from an influx of commercial deposits from weekend sales. The end-of-day vault balance tends to fall during the week, with Fridays showing the biggest decrease. The medium bank also has an increase in vault currency on Mondays, but on average there are net increases during the week and again a net drop on Fridays. The large bank shows an increase on Mondays, with a large drop on Fridays. The similar patterns on Fridays may reflect payroll withdrawals.

As we have said, vault cash also counts toward reserve requirements. The Federal Reserve does not pay interest on required reserves, but banks are able to earn credits against other services on a portion of the deposits at Federal Reserve Banks. If a bank has a clearing balance contract, there is also an incentive to maintain surplus reserves in accounts at the Fed to reduce potential overdrafts. In addition, lagged vault cash is used to satisfy reserve requirements, so that lower vault cash in

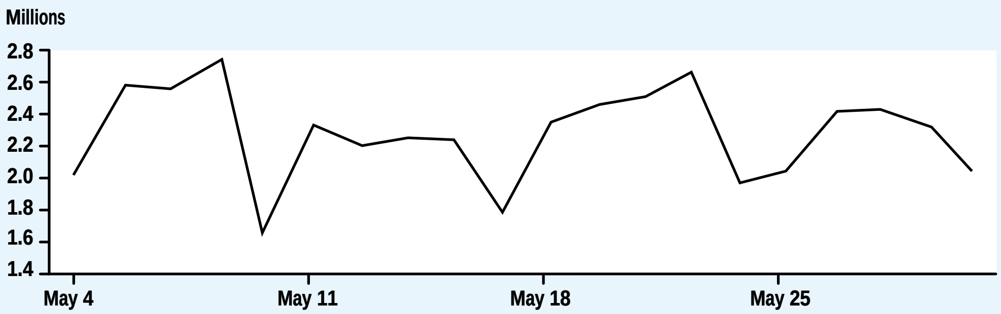
Figures 4-6

End-of-Day Vault Cash (May 1992)

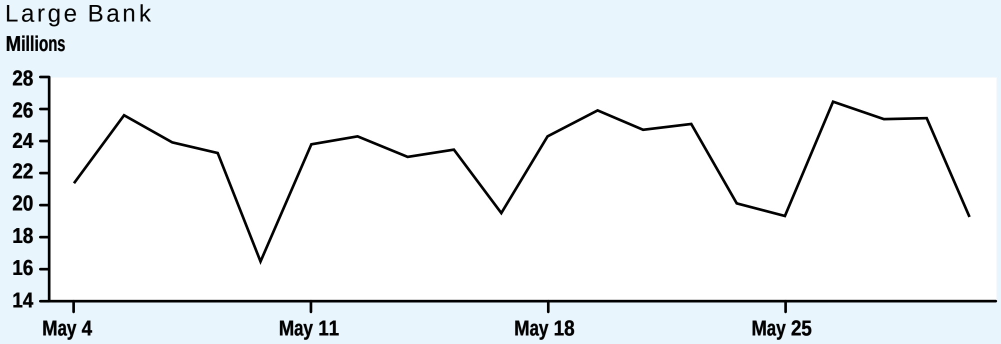
Small Bank



Medium Bank



Large Bank



the current period can be used to hedge against unexpected changes in Fed account balances. So there is some trade-off between reserves held at the Fed or vault cash. But when banks are bound, there is room for more conservative vault cash management than would otherwise occur. This seems to be the case in the Eighth District.

BANK VAULT CASH BALANCES

The end-of-day vault cash balances at all banks in the Eighth District from 1989 to 1996 were analyzed to determine the average minimum levels and replenishment points targeted by the banks. Figure 7 shows the average number of days between replenishment of vault cash for

Figure 7

Mean Days Between Vault Cash Replenishment

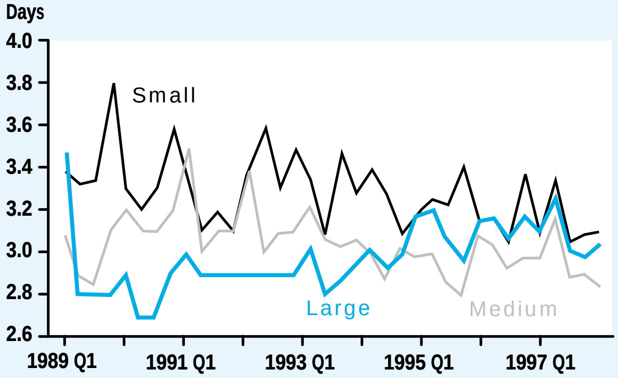


Table 2

Mean Ratio of Vault Cash to Net Transaction Deposits

	1989	1990	1991	1992	1993	1994	1995	1996	1997
Individual Banks									
Mean	4.98	4.84	4.51	4.17	3.92	4.02	4.38	4.90	4.94*
Variance	2.35	2.79	1.03	1.26	0.64	1.31	1.42	1.74	1.23*
Aggregate									
Mean	4.68	4.72	4.51	4.02	3.84	4.03	4.50	5.44	6.22*
Variance	0.12	0.13	0.11	0.10	0.07	0.10	0.11	0.24	0.44*

*1997 figures exclude small banks in the third and fourth quarters. During the last two quarters of 1997, it appears that small banks increased reclassifications, thereby lowering net transaction deposits significantly.

small, medium, and large banks in the Eighth District. As the figure shows, the typical number of days between replenishment remains relatively constant, at approximately three. Small banks show a decline in the days between replenishment since the start of the period, suggesting a possible decrease in the relative fixed cost of shipments.

Replenishment every three days would indicate a relatively volatile movement in vault cash balances (as with [*S,s*] inventory management). This volatility is lowered in the aggregate because of offsetting patterns at different banks. Table 2 shows the mean vault cash level as a percentage of net transaction deposits and the variance of these levels for individual banks and for the aggregate of all banks.[3] As expected, the volatility of total vault cash balances as a percentage of net transaction deposits is significantly less than for the average bank. So, whereas the frequent trips for currency lead to high volatility for individual banks, negative correlation among banks smooths out the volatility in the aggregate. This pattern suggests that interbank currency transfers can reduce the role of the Federal Reserve in providing currency.

Figures 8, 9, and 10 show the average upper and lower levels of vault cash as a percentage of net transaction deposits (NTD) over the period from 1989 to 1997. Observing the upper and lower bounds of the average vault cash balances, one can conclude that larger banks appear to have an average minimum balance of 4.4 percent of NTD and an average maximum of around 5.1 percent of NTD. A decline in 1997 followed a noticeable rise in the two prior years. Medium banks, on average, hold peak vault cash balances of around 4.5 percent of NTD and have minimum triggers of around 3.7 percent. Small banks fall in the lower 3 percent tranche of reserve requirements but appear to be slightly more conservative than medium banks, holding vault cash balances between 4.0 percent and 4.9 percent of NTD. This difference would suggest that most small banks are nonbound most of the time. These vault cash band widths confirm the aggregate results of Table 1. The increase in the ratio of vault currency as a percentage of NTD in large banks may also reflect the reduction in NTD due to sweeps, rather than an increase in the levels of vault cash being stored. Whether the accounts that have been moved to sweeps should be still considered transaction deposits in deciding the amount of vault cash needed is a separate issue.

To evaluate whether these levels are conservative, we need to know the distribution of intraday demands for withdrawals and deposits, the transaction costs of new shipments of currency, and the penalty that the cash manager assigns to stocking out of cash.[4] The (*S,s*) model of inventory was

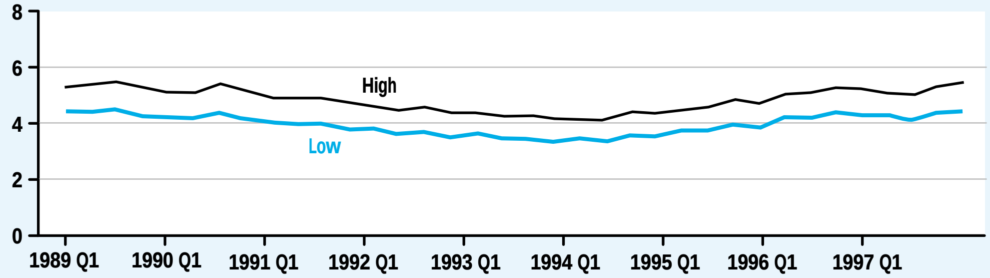
[3] The term "net transaction deposits" refers to those deposits subject to statutory reserve requirements.

[4] The vault cash figures shown here are average end-of-day figures. They do not give a good view of the fluctuations during the day. For example, large deposits from commercial customers in the evening may mask the actual minimum level that has occurred.

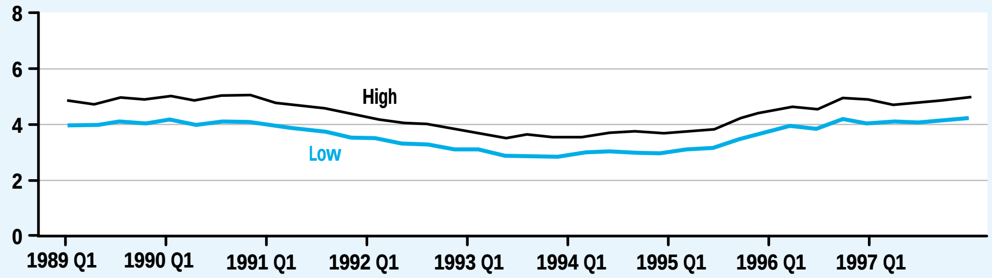
Figures 8-10

Average High and Low Vault Cash Levels
Percent of Net Transaction Deposits

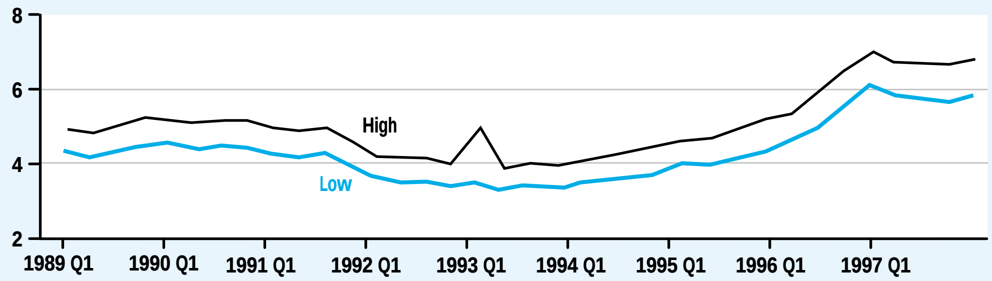
Small Banks



Medium Banks



Large Banks



formalized by Scarf (1960).[5] One way of testing the optimality of the observed levels of vault cash holding is to choose parameters for Scarf's model that suggest a band width (upper and lower limits) equivalent to the observed vault cash limits. The next section reviews the theoretical derivation of the (*S*,*s*) parameters and suggests choices of parameters to test whether banks are conservative in managing vault cash.

ARE BANKS MANAGING VAULT CASH CONSERVATIVELY?

The optimization of vault cash must trade off the cost of "stocking out" of cash against the opportunity cost of trading in the overnight markets relative to bank cash managers' level of risk aversion. The problem belongs to the class of inventory decisions. More specifically, if there are

[5] See Ross (1970) for an exposition.

fixed costs associated with obtaining currency from the Fed (for example, transportation costs), and there are zero marginal costs of purchasing currency, an (*S*,*s*) inventory rule for vault cash management may be optimal. That is, banks may establish a minimum vault cash level, *s*, below which they will replenish, and a target level, *S*, to which they will raise vault cash. When vault cash lies between these two levels, no action is taken. Although these target levels will vary with the probability distribution of customer transactions and the opportunity cost of holding cash, changes in these levels may be sticky. This inherent nonlinearity of (*S*,*s*) management can lead to difficulty in identifying interest elasticities.

The optimality of (*S*,*s*) inventory behavior hinges on the existence of nontrivial fixed adjustment costs associated with purchases (or production). In the one-period case, assuming r_t is the cost of storage and p_t is the price of the product, we can determine the penalty cost of stocking out and the cost of having inventory I_t^* at the beginning of period *t*, assuming that orders are filled instantaneously. We assume that the firm observes the inventory level at the end of the previous period and then decides the level of inventory to hold for this period. Put another way, the firm makes a decision on the appropriate level of inventory for period *t*, given the level of inventory at the end of the previous period, the expected cost of storage, and the expected cost of stocking out. See the shaded insert for a formal derivation of the limits (*S*,*s*), which appears on page 51.

The determinants of the band width (*S*,*s*) for inventory are (1) the distribution of demand (represented by the mean, μ, and the standard deviation, σ), (2) the price of the product, *p*, (3) the cost of storage, *r*, and (4) the "shape" of the cost curve (or, more directly, the shape of the marginal cost curve). The interval can be expressed as a function, *h*(.), of these parameters:

$$[S,s] = h(\mu,\sigma,p,r,a_0,a_1,a_2),$$

where a_0, a_1, and a_2 are the constant term (fixed cost), the coefficient on the linear term, and the coefficient on the quadratic term of the cost equation, respectively. Note that *h*(.) is a highly nonlinear and discontinuous function. The important parameters are the relative markup between price and marginal cost (which determines the benefit of adjusting) and the probability distribution of demand (which determines the relative cost of storage).

Because the (*S*,*s*) model is nonlinear, it is difficult to measure the effects of changes in parameters on the desired level of stock by observing contemporaneous changes. In other words, a change in (*S*,*s*) may or may not result in a contemporaneous change in inventory level, depending on the initial location of the firm's inventories within the band. For instance, if *S* were to rise as a result of a decrease in interest rates but there were sufficient inventory to delay replenishment, no concurrent move would be observed in inventory levels; in fact, inventory would fall. If firms needed to replenish at the same time that the desired upper bound increased, inventories would rise above normal. Interest rate elasticities computed under the two different initial conditions would be underestimated in one instance and overestimated in the other.

The upper bound, *S*, increases with (1) an increase in mean demand, (2) an increase in the variance of demand, (3) a decrease in the cost of storage (*r*), or (4) an increase in the markup over marginal cost (p-a_1).

Similarly, the lower bound, *s*, decreases with (1) an increase in the "quasi-fixed" cost (a_0), (2) an increase in the cost of storage (*r*), (3) an increase in the marginal cost, (4) a decrease in the price (*p*), (5) a decrease in the mean demand, or (6) an increase in the variance of demand.

Adaptation to Vault Currency

To apply the Scarf inventory model to vault currency management, we need to make some modifications.[6] In particular, the "cost" associated with obtaining currency is limited to the "quasi-fixed"

[6] Poole (1968) applies a similar inventory model to reserve management. Here, the focus is primarily on currency held in vaults, rather than total reserve management. Orr and Mellon (1961) also used this method to analyze expansion of bank credit.

Derivation of the (*S*, *s*) Inventory Rule

If demand is a random variable with known probability distribution $g(D)$, and we have an initial inventory level of I_{t-1} at the end of period $t-1$, then we can formulate the sum of the expected penalty and holding cost of purchasing sufficient quantity to have an inventory level of I_t^* at the beginning of period t as $L(I_t^*)$

$$L\left(I_t^*\right) = p_t \int_{I_t^*}^{\infty} \left(\xi - I_t^*\right) g\left(\xi\right) d\xi + r_t \int_0^{I_t^*} \left(I_t^* - \xi\right) g\left(\xi\right) d\xi. \tag{1}$$

The first term represents the expected lost revenue from stocking out, when demand exceeds inventories; it is a decreasing function of the inventory level. The second term reflects the cost of storage, or the unit storage cost times the expected excess of inventory over demand; it is an increasing function of inventory level. $L(I_t^*)$ is U-shaped or V-shaped, reflecting the sum of the downward-sloping expected cost of stocking out and the upward-sloping cost of holding inventory.

Assuming linear costs of ordering,[1] if we order $Q_t = I_t^* - I_{t-1}$, then the total costs (of ordering, stockout, and inventory) can be expressed as follows:

$$f\left(Q_t\right) + L\left(I_t^*\right) = \begin{cases} a_0 + a_1\left(I_t^* - I_{t-1}\right) + L\left(I_t^*\right), & \text{if } I_t^* > I_{t-1} \\ L(I_{t-1}) & \text{if } I_t^* = I_{t-1} \end{cases}. \tag{2}$$

Scarf (1960) shows that if we define S as the value of I_t^* that minimizes $a_1 I_t^* + L(I_t^*)$ and s as the value I_t^* that makes

$$a_1 s + L(s) = a_0 + a_1 S + L(S), \tag{3}$$

then it can be shown that the optimal policy is

$$\begin{cases} \text{if } I_{t-1} \geq s, & \text{do not order and} \\ \text{if } I_{t-1} < s, & \text{order up to } S. \end{cases} \tag{4}$$

We can further determine S from the following result:

$$G(S) = \frac{p - a_1}{p + r_t}, \tag{5}$$

where $G(S)$ is the cumulative distribution of the demand. Using the definition in Equation 3, we can then obtain the value of s.

[1] i.e., $f(Q_t) = a_0 + a_1 Q_t$

costs (transportation and administrative) associated with each trip to the Federal Reserve Bank, correspondent bank, or other source of currency shipments.[7] The "cost" of stocking out of currency is not simply the lost revenue from a transaction, it also involves lost customer confidence and future implications. This variable may be as much subjective as it is real. Since cash flow will be in both directions (in contrast to the monotonic decrease of inventory), the distribution of demand for

[7] The Fed has instituted limits on the number of free currency shipments allowed by member banks. This limit would need to be factored into the assumed cost of currency.

Table 3

Sample Parameters for Equation 6

Case	Mean (μ)	S.D.(σ)	p	r	a_0	a_1	S	s
1	0.27	3	0.1	0.02	0.0025	0	4.31	3.48
2	0.27	3	0.05	0.02	0.0025	0	3.34	2.43
3	0.27	3	0.05	0.02	0.002	0	3.34	2.52
4	0.27	3	0.02	0.02	0.002	0	2.13	1.20
5	1.0	3	0.02	0.02	0.002	0	2.44	1.54
6	0.3	5	0.10	0.02	0.0025	0	7.09	6.00
7	0.3	2	10	0.02	0.0025	0	6.41	5.96
8	0.3	2	5	0.02	0.0025	0	5.99	5.51
9	0.3	2	1	0.02	0.0025	0	4.88	4.36

vault cash is the joint distribution of demand for withdrawals of currency and the supply of currency deposits.

As an upper bound, we can assume that the cost of stocking out exceeds the storage cost by a factor of 10. From Equation 5 it can be shown that the maximum (threshold) of vault cash, S, is determined by the value that equates $G(S)$ to $p/p+r$. Thus if p is 10 times r, this fraction becomes 10/10+1, which is around 0.91. This result suggests that the maximum level of vault cash should be about two standard deviations above the average daily demand for currency. The lower limit, or trigger point, is a function of the quasi-fixed transaction cost—in this case, primarily transportation costs. If the penalty portion of the cost curve (i.e., the expected cost of running out of cash) is steep, then the minimum will be close to the maximum. In this situation, we would expect banks to frequently replenish vault cash to their maximum levels.

As a lower bound, if we assume that the cost of stocking out is equal to the cost of storage (foregone interest), then the penalty function is symmetric and $G(S) = r/(r+r)$, or 0.5. This result suggests that the maximum vault cash level would be the mean daily demand for vault cash. Again, the minimum level s would depend on the "quasi-fixed" cost, or transportation cost, in such a way that replenishment takes place when the expected cost of a shortfall exceeds the quasi-fixed cost. A choice of S as the mean daily cash demand would most likely result in at least a daily replenishment of vault cash. If the fixed (transportation) cost were small enough, the trigger would be relatively close to the maximum and would imply replenishment more than once throughout the day.

The penalty for running out of vault cash is difficult to define. If a bank can defer demand for large amounts of cash, then outliers can be eliminated from the assumed distribution of demand. Similarly, if cash can be obtained from the Federal Reserve or other banks on relatively short notice, then a smaller upper bound (S) and a narrow band width can be accommodated easily.

Calibrated Cases

The key to determining whether banks are conservatively managing vault cash is by knowing the probability distribution of the intraday deposits and withdrawals. A good model of the costs associated with stocking out is also required. Unfortunately, we have neither of these. The model above can be calibrated with specific choices for the coefficients to give the average lower and upper bounds, shown in figures 8-10. Table 3 shows nine cases with assumed parameter values for Equation 6 and the resulting upper and lower bounds. For Case 1, the parameters are chosen to give upper and lower limits equivalent to the 1993 second-quarter averages for the upper and lower limits for small banks, shown in Figure 8. These parameters are consistent with the observed average of three days between replenishment. The mean demand was chosen to produce replenishment every three days when the separation of S and s is 0.81 percent of NTD. The cost of maintaining vault balances was assumed to be approximately the overnight rate of borrowing. The penalty cost, p, and standard deviation, σ, were chosen to fix the upper limit at 4.31 percent of NTD, and the fixed cost, a_0, was chosen to fix the lower limit at 3.48. The standard deviation chosen is large, compared with the mean value of demand, and suggests a great deal of uncertainty in the withdrawal/ deposit rates. By way of comparison, Table 2 shows a standard deviation of 1.67 of the mean vault cash holdings of all Eighth District banks in 1990. This is obviously not the same as the standard deviation of daily demand, but it represents the fluctuation around mean vault cash levels for all banks.

Cases 2 through 6 show the changes to the (*S*,*s*) interval as the other parameters are changed. Case 2 shows the effect of reducing the penalty cost of running out of currency by one-half, from five times the opportunity cost of holding extra vault cash. The upper bound, *S*, falls by 1 percent of NTD to 3.34 percent of NTD. In Case 3, the fixed cost associated with currency shipments is lowered, making more frequent trips economical. This raises the trigger point, s, for replenishment. Case 4 shows the impact of assuming symmetric costs of storage and stocking out. This situation reduces the upper and lower bounds. Increasing the mean demand, as in Case 5, raises the upper and lower bounds. Case 6 shows that raising the standard deviation of demand from three to five, compared with Case 1, increases the band width to the level shown for large banks in Figure 10. Case 7 shows the minimum and maximum vault cash levels that would occur if a "more reasonable" standard deviation of demand were used, but a very high penalty factor—equivalent to 500 times the cost of storage—were applied to running short of cash. In this case, the minimum and maximum are in the same order as those recorded in Figure 10. Cases 8 and 9 show the reductions in minimum and maximum vault cash levels that would occur if the penalty assigned to running out were lowered to 250 times and 50 times the storage rate, respectively.

POLICY IMPLICATIONS

The data through 1997 appear to support the idea that banks have not been managing vault cash holdings very closely. Within the context of an (*S*,*s*) inventory model, the variance of net withdrawals and/or the penalty associated with running out would have to be very high to justify the levels of vault cash balances. When banks are bound, there is less incentive to optimize currency holdings. Because the reserve requirement levels have fallen as a result of the rate reduction and the implementation of "sweeps," nonbound banks have had more incentive to reduce surplus vault cash. Some banks have begun to use consultants to implement vault cash management strategies. As this practice becomes more prevalent, currency operations at Federal Reserve Banks could see a temporary net inflow of currency and possibly more frequent shipments. There may also be an increase in the noise in the measurement of monetary aggregates and a rise in the elasticity of bank reserves to interest rate changes. The implications for safety and soundness of depository institutions are not immediately clear, but are not expected to be large. It would appear that if banks maintain less buffer stock holdings of currency, then they may be more sensitive to increased demand for currency. However, innovations in the payments system are likely to reduce the demand for currency in the future.

In summary, the degree to which banks optimize their vault currency holdings has implications for the Fed's currency management, monetary policy effectiveness, and the safety and soundness of financial institutions. For these reasons, vault currency management should be monitored. Initial indications are that depository institutions have begun to pay closer attention to vault cash needs; the impact of this new awareness may be minimal but should be assessed, nonetheless.

REFERENCES

Anderson, Richard G., and Robert H. Rasche. "Measuring the Adjusted Monetary Base in an Era of Financial Change," this *Review* (November/December 1996), pp. 3-37.

Friedman, Milton, and Anna Jacobson Schwartz. "*A Monetary History of the United States 1867-1960*," Princeton University Press, 1963.

Gilbert, R. Alton, and Jean M. Lovati. "Bank Reserve Requirements and Their Enforcement: A Comparison Across States," this *Review* (March 1978), pp. 22-31.

Poole, William. "Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy," *Journal of Finance* (December 1968), pp. 769-91.

"Open Market Operations During 1996," *Federal Reserve Bulletin* (July 1997), pp. 565-74.

Orr, Daniel, and W. G. Mellon. "Stochastic Reserve Losses and Expansion of Bank Credit," *American Economic Review* (September 1961), pp. 614-23.

Ross, Sheldon M. *Applied Probability Models with Optimization Applications*, Dover, 1970.

Scarf, Herbert E. "The Optimality of (*s,S*) Policies for the Dynamic Inventory Problem," *Proceedings of the First Stanford Symposium on Mathematical Methods in the Social Sciences*, Stanford University Press, 1960.