http://clevelandfed.org/research/review/
Economic Review 2000 Q3

Optimal Use of Scale Economies in the Federal Reserve's Currency Infrastructure

by Paul W. Bauer, Apostolos Burnetas, Viswanath CVSA, and Gregory Reynolds

Paul W. Bauer is an economic advisor at the Federal Reserve Bank of Cleveland, Apostolos Burnetas is an associate professor of operations research at Case Western Reserve University's Weatherhead School of Management, Viswanath CVSA is an associate at McKinsey & Company, and Gregory Reynolds is an actuarial consultant at Hewitt Associates.

Introduction

The Federal Reserve has been providing paper currency since its founding in 1913. In fact, the major purpose for creating the Federal Reserve System was to furnish an "elastic currency" that would make financial crises, such as the panic of 1907, less severe. While much has changed in the financial landscape over the last 87 years, currency remains an integral part of the U.S. payments system, accounting for over 80 percent of all transactions.[1] (For a brief history of the Federal Reserve's role in currency processing, see Good and Mitchell [1999].)

For many types of transactions, currency has withstood the onslaught of checking accounts and credit cards, and will likely prevail over debit and smart cards for some types of transactions in the foreseeable future. The reasons for this are simple: Currency offers finality, anonymity, and familiarity at a reasonably low cost for small-value transactions. This sets a high hurdle for competing payment instruments. This is good news for the Treasury because the 20 billion Federal Reserve notes in circulation (with a value of $460 billion) are backed by Treasury debt. The $20 billion in annual interest payments generated by the backing of these notes is indirectly remitted to the Treasury.[2] Ultimately, the interaction of payment-instrument providers, payors, payees, and regulators will determine various instruments' market shares, but currency will probably continue to retain a significant share of transactions.

The Federal Reserve is responsible for operating its service efficiently. Even if the market were not evolving as a result of new payment instruments, the advent of nationwide branching of private depository institutions would

1 Of course, cash accounts for a much smaller share of the value exchanged in trade. Electronic payment services, such as the Clearing House Interbank Payment System (CHIPS), Fedwire, and the Automated Clearing House (ACH), now account for over 99 percent of the value exchanged in trade. See Humphrey, Pulley, and Vesala (1996) for more details.

2 Anywhere from one-half to two-thirds of U.S. currency is held overseas. Foreign holdings generally involve the higher denominations, mostly $100 bills (see Porter and Judson [1996]).

have necessitated a reexamination of the currency infrastructure.[3] This paper studies whether a redeployment of resources can lower the Federal Reserve's costs while still providing roughly the same level of service to depository institutions. More precisely, we explore the solutions to two optimization models for the Federal Reserve, identical in structure and differing only in the data fed into them. The first starts with 37 processing sites and determines how the volume should be allocated among them.[4] In other words, if we consider only the locations where the Federal Reserve already has sites, how should we allocate processing volume to minimize overall processing and shipping costs? In the second case, we start with the cities that anchor Rand McNally's 46 major trading areas (MTAs), except Honolulu.[5] This is essentially a green-field approach: If the Federal Reserve were starting from scratch, in which of the 46 largest metropolitan areas would it locate processing sites to minimize overall costs? This scenario gives us an estimate of the maximum possible cost savings from reallocating currency volume.

Like the ultimate question to life, the universe, and everything,[6] the Federal Reserve has not been given an explicit objective with respect to currency provision. Clearly, it is expected to run its currency operations in a cost-efficient manner. Furthermore, the Uniform Cash Access Policy (see footnote 3) specifies the level of service that depository institutions should receive from existing Federal Reserves sites. Beyond this, the Federal Reserve's performance objective is vague at best. Consider the question of how many processing sites the Fed should operate. More sites would mean a higher level of service to depository institutions. When a product is offered at either a lower price or higher quality, more of that product is demanded. In this case, circulating more currency would indirectly reduce the Treasury's borrowing needs. Alternatively, fewer sites would lower service levels, but might lower costs more than enough to offset any such losses to the Treasury.

The Federal Reserve generally does not address these issues directly. Before the Fourth District removed currency processing from its Pittsburgh office in 1998, the last site to be closed was the Twelfth District's Spokane office in 1938.

Because such issues are far beyond the scope of this paper, our model seeks to provide depository institutions with whatever level of service they currently receive. By adopting this constraint, we ensure that cost savings for the Federal Reserve are not obtained by making its customers worse off. Our model accomplishes this by having the Federal Reserve continue to pay for shipping currency to and from any site that is closed. This accounts for most of the social costs that such a closing would impose on third parties.[7] Whether the Federal Reserve actually picks up these costs is a policy matter that we will not consider here. The same is true of the cost implications for other Federal Reserve services remaining at these sites.[8]

In the context of production planning, distribution and logistics, mathematical models analogous to that in the present paper have been used in operations management to study problems of volume reallocation and of scale economies in processing costs. Early studies include Hanssmann and Hess (1960) and Haehling von Lanzenaur (1970) for the joint problem of production and employment planning. Recently, Thomas and McClain (1993) provide a survey of mathematical programming models in aggregate production planning, with linear, concave, or convex costs and distribution opportunities. Silver, Pyke, and Peterson (1998)

■ **3** This process began in April 1996 with the announcement of the Uniform Cash Access Policy (UCAP), designed to achieve a uniform, consistent level of cash access service across the nation in the distribution of currency. For a detailed discussion of the UCAP, see the *Federal Register,* April 25, 1996.

■ **4** After the data for this study were collected, the Pittsburgh site ceased currency processing operations, but retained paying and receiving operations. Consequently, the Federal Reserve currently operates only 36 processing sites.

■ **5** These 46 MTAs include the 37 cities with Federal Reserve processing sites (except Helena) plus 10 others. While Honolulu is also a designated MTA, no sites in Hawaii or Alaska are considered in our analysis. Note that some MTAs encompass more than one city.

■ **6** See Adams (1980).

■ **7** This is not exactly what the Federal Reserve Bank of Cleveland did when the Pittsburgh branch's high-speed sorting operation was moved to Cleveland. In this case, paying and receiving have so far been maintained in Pittsburgh. Our model assumes that these operations are also removed from any site that is closed.

■ **8** If a service is removed from a location, the building space previously allocated to it and possibly some overhead expenses must be recovered by the remaining services. This could have an adverse effect on priced services that have to recover their full economic costs through user fees.

discuss the advantages and disadvantages of using linear programming to model production problems with nonlinear cost structures. For general applications of mathematical programming methodologies in industrial, business, and economic planning, see, for example, Charnes and Cooper (1961), Hillier and Lieberman (1995), and Winston (1994).

The model we develop has applicability to any enterprise that must provide a good or service combining some activities that have scale economies with the requirement of delivering the good or service to geographically dispersed consumers. The Federal Reserve's own check processing service meets these criteria (see Bauer, Burnetas, and CVSA [forthcoming]), but so do many private and public enterprises such as joint ventures and franchises.

The key finding from the optimization of our first model is that while the Federal Reserve may be able to save almost 20 percent of its controllable costs by reallocating volume, it can obtain most of the $5 million in cost savings by reallocating processing volume without closing any processing sites. In addition, only a handful of Federal Reserve sites appear to be candidates for closing, a decision which would require examining several issues that we avoid here, such as transition costs and the impact on other Federal Reserve services. Finally, our second, green-field model suggests that the Fed's current geographic distribution of sites is very close to the optimum. The most significant departures from the status quo are that this model would not operate a site in Helena and would open sites in Phoenix and Milwaukee.

The rest of the paper is organized as follows. Section I provides an overview of the Federal Reserve's currency service. The mathematical programming model that we employ to determine the optimal allocation of processing volumes is briefly presented in section II. This model is subsequently solved using estimates for the demand for cash and the processing and shipping costs as analyzed in section III. The results are presented in section IV, together with a discussion on the robustness of the solution for a range of estimated shipping costs. Some conclusions and possible extensions to the model are discussed in section V.

I. Currency Service Overview

The Federal Reserve supplies depository institutions with currency and accepts deposits of currency from them. It also plays an important role in maintaining the quality of circulating currency by culling counterfeits and unfit notes.

Processing and distribution expenses exceed $280 million a year, and the cost to the Federal Reserve for new notes purchased from the Bureau of Engraving and Printing totals more than $400 million annually. In return, the nation's stock of 20 billion outstanding Federal Reserve notes—a total value of $460 billion—is maintained at high levels of fitness and integrity, the two components of note quality.[9]

Reserve Banks acquire new notes from the Bureau of Engraving and Printing and receive used notes from banks depositing their excess currency holdings. These paying and receiving operations require a highly secure area that separates armored car personnel from Federal Reserve employees. Security cameras cover every angle.

Deposits are received as bundles made up of 10 straps, each strap containing 100 notes. The bundles are counted manually in the receiving area, and the entire batch is cataloged and stored in the vault until it can be processed. Soon after it is deposited at a processing site, each note is counted, verified on high-speed sorting equipment, and examined by sensors that judge its fitness for circulation. The high-speed equipment then repackages fit currency into straps and bundles, which are stored in the vault until needed.

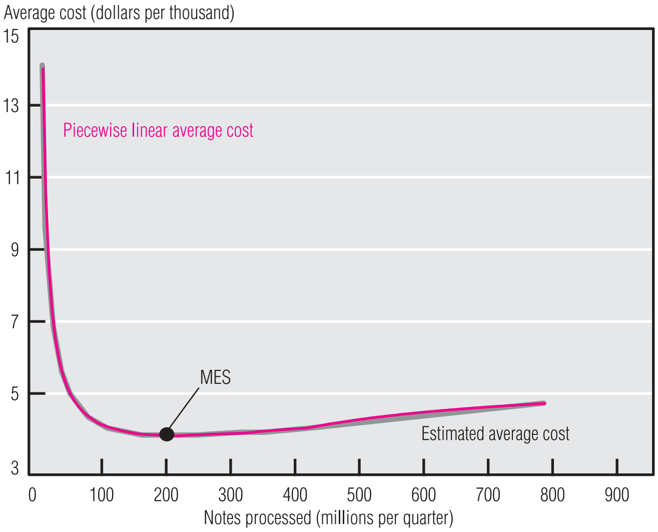
Shredders attached to the high-speed equipment destroy unfit notes. A note is deemed unfit for circulation if it is torn or has holes or if it is too soiled. Notes that are judged to be counterfeit or that cannot be read by the high-speed equipment are classified as rejects. Rejects are sent through a cancellation procedure in which operators manually examine each note and then pass it through a low-speed machine which, in conjunction with the high-speed machine, reconciles the account of the depositing bank. When a counterfeit is detected, the amount of the note is deducted from the depositing bank's reserve-account balance and the note is turned over to the Secret Service.

Currency enters circulation when a bank places an order for it; orders are filled using

9 See Bauer (1998).

FIGURE 1

Comparison of Estimated and Piecewise Linear Approximation of Average Costs



new currency or existing fit currency, depending on availability. Banks cannot specifically request new currency from the Federal Reserve.

The Depository Institutions Deregulation and Monetary Control Act (1980) does not require Reserve Banks to recover the costs of providing currency, unlike other Federal Reserve payment services (check, ACH, Fedwire, and Book Entry Securities). Generally, the only cost to banks for depositing or obtaining currency is that of transporting it to and from the Federal Reserve. The Fed's currency services are rationed according to the Uniform Cash Access Policy, designed to achieve a uniform, consistent level of cash access across Federal Reserve districts (see footnote 3). Prior to 1996, each district had its own set of policies governing currency distribution.

The Uniform Cash Access Policy limits the number of bank offices that can obtain free currency services from the Federal Reserve. A bank may designate up to 10 offices to receive one free order each week from the local Reserve Bank facility. If offices wish to deposit or order currency more frequently, they must meet certain requirements. To deposit more often, they must order more than 20 bundles in aggregate, and each order must meet the local facility's minimum threshold for each denomination requested. To order more often, offices must deposit more than 20 bundles in aggregate, and every order must meet the local facility's minimum threshold for each denomination deposited. A bank may obtain free access for more than its 10 designated offices under certain conditions: All the offices (including the designated 10) must deposit and order currency in volumes exceeding the Federal Reserve facility's high-volume threshold (generally 50 to 100 bundles) and all must meet the facility's minimum threshold for each denomination deposited or ordered. Banks that cannot meet these requirements but still wish to obtain service more frequently, or for more offices than the policy allows, may do so by paying an access fee.

II. Model Description

In this section, we briefly present the integer linear programming model we employ to determine the optimal mix of cash processing volumes and shipment schedules among processing sites that will minimize the Federal Reserve's processing and shipping costs.[10] The model is then used to explore the trade-off between scale economies in processing costs on the one hand and transportation costs on the other.

The interaction of scale economies and shipping costs is the crucial component of our model. Economists define the minimum efficient scale (MES) as the lowest level of output at which average cost reaches its minimum (see figure 1). Processing costs would be minimized if all sites operated at this level of output. To make our model computationally feasible, we employ a piecewise linear approximation to the translog cost function estimated by Bauer, Bohn, and Hancock (forthcoming). Both the translog and piecewise linear average cost functions are plotted in figure 1 to demonstrate that very little information about the estimated average cost function is lost in the piecewise approximation.[11] Note that MES is achieved at 218 million notes per quarter.

If geography—and, consequently, shipping costs—were not an important factor, then one could determine an upper bound on the number

10 While the general reader can safely skip the mathematical detail, the broad outlines of the model should be described in order to show the model's utility and limitations.

11 Bauer, Bohn, and Hancock (forthcoming) actually estimate a cost function with three outputs, fit notes, unfit notes, and shipments. For tractability, however, at each site we hold the ratio of these three outputs constant at the overall sample mean so that we only have to track processed notes (fit plus unfit notes).

of sites by dividing system-processing volume by MES. If shipping costs are low or if MES is achieved at a low level of output (that is, full-scale economies are achieved at a low level of output), then this approach would still be roughly correct. However, when MES occurs at a relatively high level of output compared to total-system processing volume, then the number of processing sites will depend on both the degree of scale economies and how expensive it is to ship currency.

We allow the model to determine whether some or all the unprocessed cash collected at one site should be shipped to other sites for processing. We also allow fit cash at a site to be shipped to other sites for distribution to the local depository institutions. The main assumptions adopted in our analysis are:

- All the locations can be used for cash processing (for some of the MTA cities, this would mean that processing facilities would have to be constructed).
- Bills are differentiated according to their denomination because shipping costs and the proportion found unfit during processing vary by denomination. The model presented in this section could differentiate between all the denominations; however, in the numerical computations described in section III we only use two types, namely, bills of low value ($1) and high value (all other denominations).
- The costs associated with shipping cash (fit or unprocessed) are proportional to the volume of shipped cash. The unit transportation cost between two sites depends on the sites chosen as well as the type of bill shipped (Insurance costs make $1 notes much cheaper to ship). The unit costs of shipping fit and unprocessed currency are the same.[12] Finally, the cost of shipping new cash to a site is directly proportional to the volume of new cash shipped, with the unit cost dependent on the destination site.
- The costs associated with currency processing are a function of the processed volume, independent of the type of bills processed. This function can generally be different for each processing site, although in the current application all sites are assumed to have the same one. This gives our approach a decidedly long-run perspective.
- There is no restriction on the amount of cash that can be stored at any site. We should note that some Federal Reserve sites are starting to face vault-capacity constraints and might have to expand vault space if processing volumes at those sites increased significantly. However, our approach takes a long-term perspective and assumes that these capacity constraints are resolved if more volume is shipped to such sites.

We use the following notation in developing our approach:

System Parameters

N = Number of cash processing sites.

b = Number of generic note types.

c_{ijk} = Unit transportation cost for currency notes of type k (processed or unprocessed) shipped from site i to site j, for $i=1,\ldots,N$, $j=1,\ldots,N$, $k=1,\ldots,b$. This parameter represents all transportation-related costs, including the cost of paying and receiving.

p_{jk} = Unit cost of new cash of type k delivered to site j, $j=1,\ldots,N$, $k=1,\ldots,b$. This parameter includes all costs due to shipping, printing, paying, and receiving new cash.

d_{ik} = Demand for currency of type k at site i.

s_{ik} = Supply of unprocessed currency of type k at site i.

u_i = Cash processing capacity at site i.

α_{ik} = Proportion of unprocessed cash at site i that is fit for circulation after processing (yield).

Decision Variables

v_{ik} = Volume of cash of type k to be processed at site i.

v_i = Total volume of cash to be processed at site i.

t_{ijk} = Volume of unprocessed cash of type k shipped from site i to site j, $i,j=1,\ldots,N$, $i\neq j$, $k=1,\ldots,b$.

m_{ijk} = Volume of fit cash of type k shipped from site i to site j (to satisfy site-j demand), $i,j=1,\ldots,N$, $k=1,\ldots,b$. Note that for $i=j$, m_{iik} denotes the amount of cash either processed at site i or sent new to site i, which is not shipped anywhere but instead is used to satisfy the demand at this site.

n_{ik} = Volume of new cash of type k sent to site i, $i=1,\ldots,N$, $k=1,\ldots,b$.

12 See Good and Mitchell (1999) for more details.

The Optimization Model

Let $f_i(v)$ denote the cost of processing cash volume v at site i.

The mathematical model developed below determines the currency volumes v_i, t_{ijk}, m_{ijk}, n_{ik} that would minimize the total processing and transportation cost, subject to shipping and sorting technology and demand-satisfaction constraints.

Minimize

$$\sum_{i=1}^{N} f_i(v_i)$$

$$+\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq 1}}^{N}\sum_{k=1}^{b} c_{ijk}t_{ijk}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq 1}}^{N}\sum_{k=1}^{b} c_{ijk}m_{ijk}$$

$$+\sum_{j=1}^{N}\sum_{k=1}^{b} p_{jk}n_{jk},$$

subject to

(a) $\sum_{j=1}^{N} m_{jik}=d_{ik} \qquad \forall i=1,\ldots,N,\ k=1,\ldots,b$

(b) $v_{ik}+\sum_{\substack{j=1\\ j\neq 1}}^{N} t_{ijk}=s_{ik}+\sum_{\substack{j=1\\ j\neq 1}}^{N} t_{jik} \qquad \forall i=1,\ldots,N,\ k=1,\ldots,b$

(c) $\alpha_{ik}\,v_{ik}+n_{ik}=\sum_{j=1}^{N} m_{ijk} \qquad \forall i=1,\ldots,N,\ k=1,\ldots,b$

(d) $v_i=\sum_{k=1}^{b} v_{ik} \qquad \forall i=1,\ldots,N$

(e) $0\leq v_i\leq u_i \qquad \forall i=1,\ldots,N.$

$m_{ijk}, t_{ijk}, v_{ik}, n_{ik}\geq 0$

We next explain briefly the motivation for these constraints.

(a) *Demand Constraints*. Cash demand at each site must be satisfied. Note that the left side of (a) is equal to the total processed cash volume available at site i, including cash sent from other sites and cash remaining at the site after processing.

(b) *Unprocessed Cash Balance*. This constraint ensures that the amount of outgoing unprocessed cash from site i (either shipped to other sites before processing, t_{ij}, $j\neq i$, or remaining at the site for local processing, v_i) is equal to the unprocessed cash coming into the site (shipped from private banks, s_i, or from other sites t_{ji}, $j\neq i$).

(c) *Fit Cash Balance*. This constraint ensures that the amount of incoming fit cash equals the outgoing fit cash at each site.

(d) The total volume of cash processed at a site comprises the volumes of different types.

(e) This last constraint ensures that processing capacity is not exceeded at any site.[13]

The model above is a nonlinear programming problem because the processing-cost functions $f_i(v)$ are generally nonlinear. To speed the computation of the solutions to our various models, we transform the original model into a mixed-integer, linear-programming problem.

The Processing Cost Function

The model will be computationally tractable as long as the processing cost functions, $f_i(v)$, can be adequately represented by a piecewise linear function as

$$f_i(v)=f_{ij}+\Theta_{ij}(v-w_{ij}),$$

$$\text{for } w_{ij}<v\leq w_{i,j+1}, j$$

$$=1,\ldots,r_i-1, \qquad (1)$$

where v is volume, i indicates the processing site, $w_{il}=0$, $w_{ir_i}=u_i$, and $f_i(0)=0$. According to this definition, the range $[0, u_i]$ of possible processing volumes at site i can be divided into r_i-1 consecutive subintervals with endpoints w_{i1}, $j=1,\ldots,r_i$, such that the processing cost is linear with a slope equal to Θ_{ij} within the subinterval $[w_{ij}, w_{i,j+1}]$. In addition, because the cost function is continuous in v for all $v>0$, it follows that $f_{ij}=f_i(w_{ij})$; therefore, equation (1) is equivalent to

$$f_i(v)=\Theta_{i1}w_{i1}+\sum_{l=2}^{j-1}\Theta_{il}(w_{i,l+1}-w_{il})+\Theta_{ij}(v-w_{ij}),$$

$$\text{for } w_{ij}<v\leq w_{i,j+1}, j=1,\ldots,r_i-1. \qquad (2)$$

■ **13** Given our long-run perspective, all processing sites are assigned the same maximum capacity. As this capacity was set far above the largest volume observed at any processing site, it does not turn out to be a binding constraint.

The slopes, Θ_{ij}, of the cost function in consecutive segments are increasing in j for any i; therefore, the cost is convex for $v > 0$.

The functional form of the processing costs proposed in equations (1) and (2) is consistent with empirical findings about economies of scale in cash processing.[14] Figure 1 illustrates how close the piecewise linear approximation is to the estimated translog.

The presence of scale economies is also clearly evident. Initially, average cost falls as a result of scale economies; however, once they are exhausted at about 218 million notes per quarter, average cost begins to increase slowly. Our piecewise linear approximation can be made arbitrarily accurate by increasing the number of subintervals r_i, but we deem nine segments to be close enough, given the trade-off between the number of segments and the speed of convergence.

We employ a piecewise linear approximation to $f(v)$ based on nine subintervals, with endpoints w_j (in thousands of notes) and costs f_j (in dollars), for $j=1,\ldots,10$ as specified in table 1. The piecewise linear approximation model for $f_i(v)$ allows us to reformulate the original nonlinear programming problem using mixed-integer linear programming. To see this, note that the piecewise linear cost can be equivalently defined by the sequence of points $\{(w_{ij}, f_{ij}),\ j=1,\ldots,r_i\}$. Therefore, for any $v \in (0,u_i]$, $f_i(v)$ can be expressed as the solution to the following linear programming problem:

$$f_i(v) = \min\ \Theta_{i1}\, z_1 + \sum_{l=2}^{r_i-1} \Theta_{il}(z_l - z_{l-1})$$

$$z_l \leq w_{i,l+1} \qquad l=1,\ldots,r_i-1$$

$$0 \leq z_1 \leq z_2 \leq \ldots \leq z_{r_i-1} = v.$$

Indeed, because Θ_{ij} is increasing in j, it can be shown that if v is in the j^{th} subinterval, that is, $w_{ij} < v \leq w_{i,j+1}$ for some j, then the optimal solution to the problem above is $z_l = w_{i,l+1}$, $l=1,\ldots,j-1$, and $z_j = z_{j+1} = \cdots = z_r = v$, with the objective function value precisely equal to the processing cost as defined in equation (2).

The above expression describing the processing cost function at each site i must now be incorporated into the original problem of cost minimization. Because $f_i(v)$ is not continuous for $v=0$—the discontinuity representing the fixed processing costs—to model the fixed cost, we define a binary variable δ_i for each site i, such that $\delta_i=1$ if $v>0$ and $\delta_i=0$ if $v=0$. The variable δ_i is associated with keeping site i open for cash processing ($\delta_i=1$) or not ($\delta_i=0$).

The resulting mixed-integer, linear-programming problem is given below:

$$\sum_{i=1}^{N} f_{i1}\delta_i + \sum_{i=1}^{N} \Theta_{i1} z_{i1} + \sum_{i=1}^{N} \sum_{l=2}^{r_i-1} \Theta_{i1}(z_{il} - z_{i,l-1})$$

$$\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k=1}^{b} c_{ijk} t_{ijk} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k=1}^{b} c_{ijk}\, m_{ijk}$$

$$+ \sum_{j=1}^{N} \sum_{k=1}^{b} p_{jk}\, n_{jk}$$

Minimize $\displaystyle\sum_{j=1}^{N} m_{jik} = d_{ik}$

$+$

$$v_{ik} + \sum_{\substack{j=1 \\ j\neq i}}^{N} t_{ijk} = s_{ik} + \sum_{\substack{j=1 \\ j\neq i}}^{N} t_{jik}$$

subject to $\displaystyle\alpha_{ik} v_{ik} + n_{ik} = \sum_{j=1}^{N} m_{ijk}$

$$v_i = \sum_{k=1}^{b} v_{ik}$$

$$v_i \leq u_i \delta_i$$

$$z_{il} \leq w_{i,l+1}$$

$$0 \leq z_{i1} \leq z_{i2} \leq \ldots \leq z_{i,r_i-1} = v_i$$

$$m_{ijk}, t_{ijk}, v_{ik}, n_{ik} \geq 0,\ \delta_i \in \{0,1\}_i\ .$$

Although our model is similar in its intent to that developed by Good and Mitchell (1999), there are several important differences. First, their model is geared to calculating the costs when various sites are closed, whereas our model can examine the case when only some volume is shipped to another site to more fully exploit scale economies. Stemming from this first difference, our model keeps track of low- and high-denomination notes because although both have the same processing costs, low-denomination bills are cheaper to ship, mainly because of lower insurance costs. Another significant difference is that the processing-cost function employed by Good and Mitchell does not allow for scale diseconomies. While they start with our processing-cost function, they assume that variable unit costs remain constant once minimum efficient scale is achieved. Finally, they consider some indirect costs that we ignore. For example, they reduce protection costs from sites that are closed, whereas we do not. While these are all potentially significant differences,

■ **14** See Bauer, Bohn, and Hancock (forthcoming).

TABLE 1

Endpoints of Cost-Function Approximation

Subinterval number j	Subinterval starting point w_j (thousands of notes)	Cost f_j (dollars)
1	0	100,166
2	8,913	125,687
3	28,000	183,207
4	48,000	245,717
5	98,000	406,406
6	158,000	612,088
7	218,000	833,646
8	318,000	1,238,865
9	418,000	1,688,538
10	788,000	3,726,716

SOURCES: Authors' calculations; and Bauer, Bohn, and Hancock (forthcoming).

the next section will show that, given the locations of sites, volume, scale economies, and shipping costs actually observed, the two models come up with very similar sets of results.

Some caveats should be kept firmly in mind. First, we intend to model long-run operational costs; consequently, we do not consider any of the transition costs that would surely be incurred if a site were to be opened or closed. Our results can only suggest situations that require additional study. Before actually closing a processing site, a more complete cost-benefit analysis should be performed.

Second, although our estimates of transportation costs are the best available, they are still just estimates. Without actually performing the contemplated shipments, it is unlikely that our estimates of shipping costs can be improved. In light of this, we tested the sensitivity of our results by solving the model with transportation costs 10 percent higher and 10 percent lower than our best estimates.

A third potential weakness is that the cost function estimated by Bauer, Bohn, and Hancock (forthcoming) used data from a sample that included sites with both old and new (high-speed) currency-sorting machines.[15] When a large enough sample has been collected from the new sorters (which will require the further passage of time), the cost functions could be re-estimated and our models could be reoptimized.

On a related issue, the cost functions estimated by Bauer, Bohn, and Hancock (forthcoming) incorporate site-specific environmental variables and allow for varying levels of cost efficiency. Including these factors would result in cost-function variations across processing sites, which could affect our results. On the other hand, the sources of these varying productivity levels can be mitigated over time (by encouraging underperforming sites to adopt best-practice techniques), and our model is a long-run model at heart, so an argument can be made for omitting potentially transient factors. However, one short-run factor that we may want to include in future work is vault capacity, which varies significantly across processing sites. Limited vault capacity could preclude some reallocations.

Finally, while most of the cost savings can be obtained by reallocating volumes without closing any sites, a different objective function (for example, one that specifies either higher or lower service levels) might suggest more fundamental changes. Solving our model with this alternative performance objective could provide additional insight into such important policy questions.

III. Analysis

Two sets of data are used for our computations, hereafter referred to as the Current Processing Sites (CPS) and the Major Trading Areas (MTA) data sets, respectively. Except as noted below, all the data comes from the Federal Reserve's Planning and Control System quarterly expense reports. The two sets differ in the number of sites considered cash processing centers and distribution depots and in essence represent alternative ways of allocating the same national data.

The CPS set comprises the existing 37 cash-processing sites of the Federal Reserve System (see table 2 for a list of these sites). By optimizing the model with these data, we will explore whether reallocating volume among the existing Federal Reserve sites can lower overall costs. Alternatively, the 46-site MTA set includes the entire 37 CPS set, except Helena, plus 10 others.[16] Analysis of this set explores where currency-processing sites would be located if the Federal Reserve were to start from scratch, adopting a green-field approach.

15 All of the older sorting machines are now retired.

16 Good and Mitchell (1999) provide details on how demand for currency was derived for the 46 MTA sites.

TABLE 2

Demand for Unprocessed Cash[a]

37-site case (Current Processing Sites)			46-site case (Major Trading Areas)		
	Low-value notes	High-value notes		Low-value notes	High-value notes
Atlanta	49,722	100,952	Atlanta	59,299	118,597
Baltimore	60,628	123,094	Baltimore–Washington	69,977	139,955
Birmingham	20,322	41,260	Birmingham	18,938	37,876
Boston	143,476	291,300	Boston–Providence	110,942	221,884
Buffalo	39,483	80,163	Buffalo–Rochester	41,118	82,236
Charlotte	89,775	182,271	Charlotte–Greensboro–Raleigh	87,348	174,695
Chicago	176,208	357,757	Chicago	101,817	203,633
Cincinnati	36,244	73,586	Cincinnati–Dayton	35,337	70,674
Cleveland	35,025	71,112	Cleveland	23,056	46,111
			Columbus	13,537	27,075
Dallas	43,006	87,316	Dallas–Fort Worth	54,031	108,061
Denver	33,183	67,372	Denver	26,122	52,244
			Des Moines–Quad Cities	13,032	26,065
Detroit	63,652	129,232	Detroit	71,285	142,569
El Paso	11,072	22,480	El Paso–Albuquerque	16,588	33,175
Helena	3,830	7,776			
Houston	33,771	68,566	Houston	35,299	70,599
			Indianapolis	21,014	42,028
Jacksonville	36,360	73,823	Jacksonville	10,702	21,404
Kansas City	16,690	33,886	Kansas City	12,821	25,641
			Knoxville	9,200	18,399
Little Rock	16,054	32,594	Little Rock	12,942	25,884
Los Angeles	196,467	398,887	Los Angeles–San Diego	165,664	331,328
Louisville	18,028	36,603	Louisville–Lexington–Evansville	22,388	44,776
Memphis	17,510	35,551	Memphis–Jackson	21,627	43,255
Miami	40,520	82,268	Miami–Fort Lauderdale	40,929	81,859
			Milwaukee	38,501	77,002
Minneapolis	39,139	79,464	Minneapolis–St. Paul	35,137	70,274
Nashville	19,129	38,837	Nashville	9,865	19,730
New Orleans	37,930	77,009	New Orleans–Baton Rouge	31,729	63,458
New York	293,172	595,227	New York	365,755	731,511
Oklahoma City	18,861	38,293	Oklahoma City	11,337	22,674
Omaha	9,089	18,453	Omaha	7,265	14,529
Philadelphia	83,083	168,683	Philadelphia	66,213	132,427
			Phoenix	28,303	56,605
Pittsburgh	32,025	65,021	Pittsburgh	28,475	56,949
Portland	16,379	33,255	Portland	14,476	28,952
Richmond	50,720	102,976	Richmond–Norfolk	34,777	69,554
St. Louis	25,760	52,300	St. Louis	26,865	53,730
Salt Lake City	12,139	24,646	Salt Lake City	12,295	24,591
San Antonio	25,949	52,685	San Antonio	19,891	39,782
San Francisco	103,361	209,854	San Francisco–Oakland–San Jose	95,631	191,263
Seattle	35,207	71,480	Seattle	31,097	62,194
			Spokane–Billings	9,248	18,495
			Tampa–St. Petersburg–Orlando	34,322	68,643
			Tulsa	6,664	13,328
			Wichita	5,735	11,469
Totals	**1,982,969**	**4,026,028**		**2,002,859**	**4,005,714**

a. In thousands of notes.

SOURCES: Good and Mitchell (1999); and Planning and Control System Expense Report, 1996.

TABLE 3

Supply of Unprocessed Cash[a]

37-site case (Current Processing Sites)			46-site case (Major Trading Areas)		
	Low-value notes	High-value notes		Low-value notes	High-value notes
Atlanta	46,522	94,453	Atlanta	55,481	110,962
Baltimore	58,671	119,120	Baltimore–Washington	67,718	135,436
Birmingham	17,181	34,882	Birmingham	16,011	32,021
Boston	132,935	269,899	Boston–Providence	102,791	205,583
Buffalo	38,908	78,995	Buffalo–Rochester	40,519	81,039
Charlotte	84,448	171,456	Charlotte–Greensboro–Raleigh	82,165	164,330
Chicago	162,057	329,024	Chicago	93,639	187,279
Cincinnati	29,001	58,881	Cincinnati–Dayton	28,276	56,551
Cleveland	37,313	75,757	Cleveland	24,561	49,123
			Columbus	12,222	24,443
Dallas	37,696	76,534	Dallas–Fort Worth	47,359	94,717
Denver	32,445	65,873	Denver	25,541	51,081
			Des Moines–Quad Cities	11,156	22,312
Detroit	53,815	109,261	Detroit	60,268	120,537
El Paso	14,484	29,407	El Paso–Albuquerque	21,698	43,397
Helena	3,888	7,894			
Houston	30,770	62,473	Houston	32,162	64,324
			Indianapolis	17,995	35,990
Jacksonville	41,621	84,503	Jacksonville	12,250	24,501
Kansas City	14,855	30,160	Kansas City	11,411	22,822
			Knoxville	9,705	19,411
Little Rock	15,731	31,938	Little Rock	12,682	25,363
Los Angeles	228,842	464,618	Los Angeles–San Diego	192,963	385,927
Louisville	16,649	33,802	Louisville–Lexington–Evansville	20,675	41,349
Memphis	17,203	34,927	Memphis–Jackson	21,247	42,495
Miami	58,477	118,727	Miami–Fort Lauderdale	59,068	118,136
			Milwaukee	35,411	70,823
Minneapolis	37,833	76,813	Minneapolis–St. Paul	33,964	67,929
Nashville	20,182	40,975	Nashville	10,408	20,816
New Orleans	40,774	82,784	New Orleans–Baton Rouge	34,109	68,217
New York	243,636	494,655	New York	303,956	607,911
Oklahoma City	17,975	36,496	Oklahoma City	10,805	21,610
Omaha	7,781	15,799	Omaha	6,220	12,439
Philadelphia	88,657	180,001	Philadelphia	70,656	141,312
			Phoenix	37,030	74,059
Pittsburgh	27,342	55,512	Pittsburgh	24,310	48,621
Portland	14,425	29,287	Portland	12,749	25,497
Richmond	48,119	97,697	Richmond–Norfolk	32,994	65,988
St. Louis	24,514	49,770	St. Louis	25,565	51,131
Salt Lake City	11,274	22,890	Salt Lake City	11,419	22,839
San Antonio	34,621	70,290	San Antonio	26,538	53,076
San Francisco	105,750	214,706	San Francisco–Oakland–San Jose	97,842	195,685
Seattle	33,015	67,030	Seattle	29,161	58,321
			Spokane–Billings	9,389	18,777
			Tampa–St. Petersburg–Orlando	45,729	91,458
			Tulsa	6,357	12,714
			Wichita	5,319	10,637
Totals	**1,929,408**	**3,917,284**		**1,949,494**	**3,898,989**

a. In thousands of notes.

SOURCES: Good and Mitchell (1999); Planning and Control System Expense Report, 1996; and authors' calculations.

In both data sets, the currency notes come in two varieties, low-value ($1) notes and high-value notes (all others).[17] The two types of notes differ in unit transportation cost as well as in the yield of the cash-processing operation. The level of security required for high-value shipments is much stricter; after all, at least five times the value is being shipped for a given number of notes. On the basis of interviews with cash personnel, we assume that a bundle of $1 bills can be shipped at one-tenth the cost of a bundle of high-value notes. The other difference between the two note types is that the processing yield is lower for $1 bills, most likely because they are employed in more transactions and receive much more severe abuse than other denominations.

We determined that the differences in transportation costs and yields for notes of $5 and higher are not significant enough to justify further differentiation. The model could be modified to distinguish each denomination, but this addition would add little to our analysis while greatly lengthening the amount of computer time required to obtain a solution.

The data on the demand for fit cash and the supply of unprocessed cash in each of the 37 CPSs are derived from average quarterly values for 1997.[18] Alternatively, for the MTA cities we rely on estimates from Good and Mitchell (1999). Low-value notes represent one-third of the total of unprocessed cash and high-value notes represent two-thirds. This ratio is based on average processing volumes observed across the Federal Reserve. The demand and supply volumes are presented in tables 2 and 3, for the CPS and MTA data sets.[19]

Although the Bureau of Engraving and Printing's average printing cost per thousand notes and the total cost of shipping new notes to Federal Reserve processing sites are known, the cost of shipping them to a particular site is not. Consequently, we take the cost of new notes to be the sum of the average cost of printing new notes plus the average cost per note of shipping currency (Bureau-to-Federal Reserve shipping costs/number of new notes). More formally, the cost of new cash delivered to site j is $p_{ij} = \$41$ per 1,000 notes, for $i = 1,2$, $j = 1,\ldots,N$.[20] The percentage yield of the cash-processing operations is equal to $\alpha_{1j} = 60$ and $\alpha_{2j} = 70$ for all sites.

The unit shipping cost between any two sites is based on estimates from Good and Mitchell (1999). Transportation costs increase with both volume and distance shipped. They also take into account the cost difference between low- and high-value notes.

IV. Results

We solve the optimization model using three separate assumptions about transportation costs to determine the sensitivity of our results.[21] After solving the model with our best estimates of transportation costs, we also estimated the model with the unit transportation costs uniformly lower (90 percent) and higher (110 percent). These scenarios are referred to as the low-cost and high-cost cases.

The results are summarized in table 4 for the 37-site (CPS) and in table 5 for the 46-site (MTA) data sets. The tables include information on the overall cost for the three cost scenarios, as well as a control case, where no shipments are allowed between processing sites, and the case where shipments are allowed but no sites are permitted to close. The no-shipment case corresponds to the state of affairs in cash processing that was current at the time the data were collected and serves as the basis of comparison for estimating cost savings through volume reallocation.[22] The information in tables 4 and 5 includes only sites that are closed under at least one scenario. Sites not present in these tables remain open for cash processing in all cases. The tables also show how costs break down into transportation, processing, and new-cash components in the

■ **17** The proportion of demand for these two types of notes is assumed to be the same for all sites and is set to their nationwide averages.

■ **18** We do not study the possible complications of seasonal fluctuations in the demand for currency across the various locations.

■ **19** The total volumes between the two data sets differ because of assumptions made by Good and Mitchell (1999).

■ **20** Allowing for differential shipping costs from the Bureau of Printing and Engraving to the various processing sites would give sites closer to Bureau's sites in Washington, D.C., and Fort Worth, Texas, an advantage over those located farther away. We determined that refining this aspect of the model was not a high priority at this time.

■ **21** The optimal solutions for the two data sets and the various scenarios are obtained by employing CPLEX optimization software.

■ **22** By "no-shipment case" we mean that the model has not added any shipments, not that no shipments are made. Recall that the model takes the existing configuration as given and so starts with shipments of about 1 billion unprocessed notes (mostly $1 bills).

TABLE 4

Optimal Solution Summary for Current Processing Sites Data Set (dollars)

	Low cost (90%)	Reference cost (100%)	High cost (110%)	All sites open	No cash shipments
El Paso	OPEN	CLOSED	OPEN		
Helena	CLOSED	CLOSED	CLOSED		
Kansas City	OPEN	CLOSED	OPEN		
Little Rock	CLOSED	CLOSED	CLOSED		
Louisville	CLOSED	CLOSED	CLOSED		
Oklahoma City	CLOSED	OPEN	OPEN		
Omaha	CLOSED	CLOSED	CLOSED		
Portland	CLOSED	CLOSED	CLOSED		
Salt Lake City	CLOSED	OPEN	CLOSED		
Low-value note	260,744	282,150	306,728	285,295	0
High-value note	386,972	435,075	410,316	163,822	0
Total transportation cost	**647,717**	**717,227**	**717,044**	**449,117**	**0**
Variable cost	19,825,154	19,826,972	19,795,879	19,563,818	25,070,764
Fixed cost	3,004,980	3,004,980	3,105,146	3,706,142	3,706,142
Total processing cost	**22,830,134**	**22,831,952**	**22,901,026**	**23,269,960**	**28,776,906**
Controllable costs	**23,477,851**	**23,549,179**	**23,618,070**	**23,719,077**	**28,776,906**
New cash cost	**75,236,786**	**75,236,786**	**75,236,786**	**75,236,786**	**75,236,786**
Total cost	**98,714,637**	**98,785,965**	**98,854,856**	**98,955,863**	**104,013,692**

SOURCE: Authors' calculations.

various cases. The controllable-cost figures refer to the sum of transportation and processing costs, because only these two components of total cost can be affected by reallocating cash volumes among sites. In contrast, the cost of new cash is not controllable. The unit cost of new cash is assumed to be the same for every site, and the total amount required is determined by the demand, supply, and yield figures and is independent of possible reallocations. Therefore, for optimization purposes, the cost of new cash can be considered a fixed cost of the currency operation.

A first observation from tables 4 and 5 is that transportation and processing costs, as well as cost savings resulting from the transportation option, are very similar between the 37- and the 46-site data sets. Although the subsequent discussion concentrates on the 37-site scenario, it applies to both.

A comparison of the controllable costs (total cost less the cost of acquiring new currency) in the reference-cost and no-shipments columns in table 4 reveals that allowing for shipment of cash between processing sites results in total cost savings of approximately \$5.2 million per quarter. This corresponds to savings in controllable costs of nearly 18 percent. When the total cost is considered by including the fixed cost of new cash, the savings are approximately 5 percent. Specifically, by spending an additional \$717,000 in transportation per quarter, a more efficient allocation of processing volumes can be achieved by exploiting scale economies more fully at some sites while avoiding scale diseconomies at others.

Comparing the last two columns of table 4 also demonstrates that the major part of these savings (approximately \$5 million) can be realized merely by allowing for cash shipments between sites without closing any. Relaxing the no-closure constraint yields additional savings of only about \$200,000. Thus, reallocation of cash volume through cash shipments seems to be the critical factor in improving the system's efficiency, whereas closing sites has a much smaller effect.

As the unit shipping costs rise from 90 percent to 110 percent of the reference case, costs increase for both transportation and processing.

TABLE 5

Optimal Solution Summary for Metropolitan Trading Areas Data Set (dollars)

	Low cost (90%)	Reference cost (100%)	High cost (110%)	All sites open	No cash shipments
Birmingham	CLOSED	CLOSED	OPEN		
Columbus	CLOSED	CLOSED	OPEN		
Des Moines–Quad Cities	CLOSED	CLOSED	CLOSED		
Indianapolis	CLOSED	OPEN	OPEN		
Jacksonville	CLOSED	CLOSED	OPEN		
Knoxville	CLOSED	CLOSED	OPEN		
Little Rock	CLOSED	CLOSED	CLOSED		
Nashville	CLOSED	CLOSED	OPEN		
Oklahoma City	CLOSED	CLOSED	CLOSED		
Omaha	CLOSED	CLOSED	CLOSED		
Portland	CLOSED	CLOSED	OPEN		
Salt Lake City	CLOSED	CLOSED	OPEN		
Spokane–Billings	CLOSED	CLOSED	CLOSED		
Tulsa	CLOSED	CLOSED	CLOSED		
Wichita	CLOSED	CLOSED	CLOSED		
Low-value note	245,753	255,839	276,592	302,860	0
High-value note	784,430	816,519	586,122	334,037	0
Total transportation cost	**1,030,183**	**1,072,358**	**862,714**	**636,897**	**0**
Variable cost	19,587,647	19,553,963	19,299,133	19,098,687	23,767,247
Fixed cost	3,105,146	3,205,312	3,903,474	4,607,636	4,607,636
Total processing cost	**22,692,793**	**22,759,275**	**23,202,607**	**23,706,323**	**28,374,883**
Controllable costs	**23,722,976**	**23,831,633**	**24,065,321**	**24,343,220**	**28,374,883**
New cash cost	**76,088,222**	**76,088,222**	**76,088,222**	**76,088,222**	**76,088,222**
Total cost	**98,811,198**	**99,919,855**	**100,153,543**	**100,431,442**	**104,463,105**

SOURCE: Authors' calculations.

This is expected, because as transportation costs rise, fewer notes are shipped, resulting in smaller cost savings from exploiting scale economies in currency processing. However, the increase in controllable costs is about 1 percent, which indicates that the optimal cost is fairly robust with respect to shipping-cost variations.

Several observations can be made from table 4 regarding the behavior of site closings as a function of shipping costs. Consider, for example, the processing site in Oklahoma City. It is optimal for this site to be closed in the low-shipping-cost case and open in the other two cases. This makes sense intuitively, as higher shipping costs tend to lower the volume of cash shipped, leading to more sites remaining open. On the other hand, for Salt Lake City, moving from normal to high shipping costs results in closing the site. This observation is counterintuitive when considered in isolation. However, the model's objective is to minimize the Federal Reserve Banks' costs, and the volume can be more cheaply handled at a combination of other sites under this cost configuration (for instance, much of the volume goes to Kansas City).

Most of the cash that gets shipped consists of low-value notes; relatively few high-value notes are shipped. The more expensive shipping costs for high-value notes appear to prohibit reallocating the processing of these notes given the relatively small cost savings from further exploiting scale economies. The lower shipping costs for low-value notes makes the transportation option more viable.

Only five sites are closed under all three levels of transportation costs. Four others might also be candidates for closing. Of course, this study does not examine all the factors that must be considered before any sites are closed. For instance, the study takes a long-run approach and assumes that all inputs adjust fully to the new processing volumes. This means that currency sorters are reallocated and vault space is constructed if necessary. Factoring these additional costs into a present-value analysis of the site-closing decision may reveal that closing, or opening, a particular site is too costly because transition costs are a friction that make change less likely. Also, at any sites that were closed the impact on the cost of providing other Federal Reserve services would have to be considered. Because it would be costly to reopen processing sites, there is an option value for retaining them.

Lastly, as mentioned above, analogous conclusions can be made from the MTA data set. However, this case does have a number of interesting points. First, processing sites in Phoenix and Milwaukee (not currently Fed sites) remain open under all three transportation-cost scenarios.[23] Second, in a handful of cases, a nearby city is preferred to an existing Federal Reserve site. For example, Tampa is preferred to Jacksonville in the MTA solution. Given the small cost advantage of the alternative configuration, the transition cost of relocating processing sites makes any such moves impractical.

Whereas Phoenix and Milwaukee remain open under all three transportation-cost scenarios, eight other sites added by the MTA data set do not.[24] Consequently, although there may be some opportunities for additional cost savings, it appears that the founders of the Federal Reserve System in 1913 did a remarkably good job of selecting processing sites, which continue to satisfy currency processing needs nearly a hundred years later.[25]

V. Summary

We set out to construct a model that we could use to determine the least-cost configuration of Federal Reserve currency processing sites given the trade-off between processing scale economies and transportation costs. We have several robust results. First, the Federal Reserve may be able to save up to about 20 percent of controllable costs by reallocating processing volumes, a bit more than the 10 percent predicted by Good and Mitchell (1999). This difference is probably because Good and Mitchell do not get any cost savings from avoiding scale diseconomies.

Second, most of these cost savings can be achieved without closing any processing sites by shipping mostly low-denomination bills from sites with scale diseconomies to sites with scale economies. This unexpected result cannot be confirmed by Good and Mitchell (1999) because their model was not set up to examine this question.

Another important result is that only a few processing sites appear to be candidates for closing. Among current processing sites, only nine warrant further study to determine whether their processing operations should be reallocated. In the green-field simulation, only 15 MTA sites (8 of them Federal Reserve sites) do not appear to be good choices for processing sites. Good and Mitchell's MTA optimum has 34 processing sites versus our 31. This appears to be another manifestation of different assumptions about the existence of scale diseconomies for larger sites leading to slightly different results.

Finally, even when we adopt a green-field approach and search for the optimal allocation of processing volume among the 46 MTAs, the current Federal Reserve sites generally remain open. Intriguingly, Phoenix and Milwaukee are the only added sites that remain open under all three shipping-cost assumptions. Alternatively, Good and Mitchell's model suggests that a site in Spokane would be viable.

As discussed earlier, some caveats apply to our results. Transition costs are neglected, shipping costs are uncertain, cost function estimates are based on an evolving technology, and finally, cost minimization may not be the sole performance objective of the Federal Reserve. We have tried to allow for these

■ **23** In July 1999, the Federal Reserve Bank of San Francisco announced that it had signed a contract to purchase a Phoenix site for a future cash operations center, which is scheduled to begin operating in September 2001.

■ **24** Because our cost function is based on a translog approximation to the underlying true functional form, it may overstate the diseconomies of scale once MES is achieved. If so, Milwaukee's volume would most likely be sent to Chicago for processing.

■ **25** While cities with currency-processing sites have received some boost to their economic vitality because depository institutions located there would incur lower costs in shipping currency between their branches and the currency depot, this endogenous effect is likely to be small.

shortcomings in various ways and feel that our qualitative results are robust.

References

Adams, Douglas. *The Hitchhiker's Guide to the Galaxy*. New York: Harmony Books, 1980.

Bauer, Paul W. "Currency: Time for Change?" Federal Reserve Bank of Cleveland, *Economic Commentary,* October 1, 1998.

Bauer, Paul W., James Bohn, and Diana Hancock. "Scale Economy, Technological Change, and Cost Efficiency of Federal Reserve Currency Processing," Federal Reserve Bank of Cleveland, *Economic Review,* forthcoming.

Bauer, Paul W., Apostolos Burnetas, and Viswanath CVSA. "Employing Scale Economies in Federal Reserve Check Processing Operations," Federal Reserve Bank of Cleveland, Working Paper, forthcoming.

Charnes, Abraham, and William W. Cooper. *Management Models and Industrial Applications of Linear Programming,* vols. I and II. New York: John Wiley & Sons, 1961.

Good, Barbara, and Harvey Mitchell. "Federal Reserve Banks and Branches Location Study," Federal Reserve Bank of Dallas, unpublished manuscript, 1999.

Haehling von Lanzenaur, Christoph. "Production and Employment Scheduling in Multistage Production Systems," *Naval Research Logistics,* vol. 17, no. 2 (1970), pp. 193–8.

Hanssmann, F., and W. Hess. "A Linear Programming Approach to Production and Employment Scheduling," *Management Technology,* vol. 1, no. 1 (1960), pp. 46–51.

Hillier, Frederick S., and Gerald J. Lieberman. *Introduction to Operations Research*. New York: McGraw-Hill, 1995.

Humphrey, David B., Lawrence B. Pulley, and Jukka M. Vesala. "Cash, Paper, and Electronic Payments: A Cross-Country Analysis," *Journal of Money, Credit, and Banking,* vol. 28, no. 4, part 2 (November 1996), pp. 914–39.

Porter, Richard D., and Ruth A. Judson. "The Location of U.S. Currency: How Much Is Abroad?" Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin,* (October 1996), pp. 883–903.

Silver, Edward A., David F. Pyke, and Rein Peterson. *Inventory Management and Production Planning and Scheduling*. New York: John Wiley & Sons, 1998.

Thomas, Joseph, and John O. McClain. "An Overview of Production Planning," in S. Graves, A. Rinnooy Kan, and P. Zipkin, eds., *Logistics of Production and Inventory,* vol. 4, chapter 7. Amsterdam: Elsevier, 1993.

Winston, Wayne L. *Operations Research: Applications and Algorithms*. Belmont, Calif.: Duxbury Press, 1994.