Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Banking & Finance 28 (2004) 2175–2190
Journal of BANKING & FINANCE
www.elsevier.com/locate/econbase

New evidence on the Fed's productivity in providing payments services ☆

R. Alton Gilbert [a], David C. Wheelock [a,*], Paul W. Wilson [b]

[a] Research Department, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, USA
[b] Department of Economics, University of Texas, Austin, TX 78712, USA

Received 11 April 2003; accepted 28 August 2003
Available online 18 December 2003

Abstract

As the dominant provider of payments services, the efficiency with which the Federal Reserve provides such services is an important public policy issue. This paper examines the productivity of Federal Reserve check-processing offices during 1980–1999 using non-parametric estimation methods and newly developed methods for non-parametric inference and hypothesis testing. The results support prior studies that found little initial improvement in the Fed's efficiency with the imposition of pricing for Federal Reserve services in 1982. However, we find that median productivity improved substantially during the 1990s, and the dispersion of productivity across Fed offices declined.

JEL classification: C14; G21; L11
Keywords: Payments system; Check processing; DEA

1. Introduction

The Federal Reserve is the largest provider of payments services in the world, offering electronic funds transfer, automatic clearinghouse (ACH), check clearing,

☆ This research was conducted while Wilson was a visiting scholar in the Research Department of the Federal Reserve Bank of St. Louis. We thank Paul Bauer and two anonymous referees for comments on an earlier version of this paper and Heidi Beyer for research assistance. The views expressed in this paper do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.
* Corresponding author. Tel.: +1-314-444-8570; fax: +1-314-444-8731.
E-mail address: wheelock@stls.frb.org (D.C. Wheelock).

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2003.08.006

and cash services. Paper check volumes have fallen precipitously in recent years as electronic payments media ranging from debit cards to ACH have become popular. Despite inroads by electronic funds transfer and ACH, however, the processing of paper checks still accounts for some 80% of all revenue and cost for Fed payments operations. The Federal Reserve processes approximately one-half of all checks deposited with US banks other than those on which the checks are drawn (Federal Reserve System, 2002). In 2001, the Fed processed 16.9 billion commercial checks on which it collected $765 million in revenue and entailed operating expenses of $684 million (Board of Governors of the Federal Reserve System, 2001).

The Monetary Control Act of 1980 requires the Federal Reserve to recover its costs of providing payments services plus a "private sector adjustment factor" that reflects estimates of the taxes that a private firm would pay, and a return on investment for its shareholders. By requiring the Fed to price its services in this manner, this provision of the Monetary Control Act sought to use market discipline to improve the efficiency with which Fed offices provide payments services.

Recently, the Federal Reserve has reconsidered and reaffirmed its role as a provider of retail payments services (Rivlin et al., 1998). One objective the Fed set for continuing to provide payments services was to encourage greater efficiency of the payments system. Evidence that Fed check facilities waste resources in providing payments services would indicate that the legal framework for pricing services imposed by the Monetary Control Act has not had its intended effect on the efficiency of Fed payments operations. Further it would indicate that the Fed could pursue greater efficiency in the operation of the payments system by focusing on the efficiency of its own payments facilities.

In February 2003, the Federal Reserve announced a reorganization of its check processing operations with improved efficiency its stated objective (Federal Reserve System, 2003). Responding to falling check volume, the Fed concluded that it must reduce the number of System offices that process checks in order to comply with the cost-recovery requirements of the Monetary Control Act. Pending legislation that would facilitate final settlement of payments made by check electronically without the physical presentment of checks to the paying bank could pose a new challenge for the Fed by further reducing the demand for its check clearing services that involve moving the original paper checks from collecting banks to paying banks.[1]

Previous studies of the efficiency with which the Fed provides check clearing services found little or no evidence of efficiency gains with the advent of pricing in the early 1980s. The increased availability of data with the passage of time, however, as well as recent advances in econometric methodology, provide an opportunity to gain insights that were not possible in earlier studies of Fed payments services. Further, the recent and projected decline in check volume and its implications for the Fed's ability to recover its costs of providing check processing services makes such a study timely.

This paper examines the productivity and technical efficiency of Fed check-processing offices using a non-parametric distance function estimator and newly-devel-

[1] The text of the Check Truncation Act as prepared by the Federal Reserve is available at http://www.federalreserve.gov/paymentssystems/truncation/ctact.htm.

oped methods for non-parametric inference and hypothesis testing. We perform tests of several model restrictions, including constant returns to scale, the appropriate number of outputs, and whether the distribution of inefficiency is independent of output levels and the mix of inputs. We find clear evidence that Federal Reserve check processing facilities have become more productive over time, consistent with the goals of the Monetary Control Act of 1980, but that substantial improvement did not occur until the late 1980s. Our estimates also indicate that, on average, Fed offices could feasibly reduce input usage by about 30% without reducing output. This level of inefficiency is similar to what others have found for commercial banks and other financial service firms. [2] Unlike previous studies, however, we present evidence of the statistical precision of our productivity estimates for individual Fed offices. Finally, our estimates indicate that the technology of check processing is characterized by variable returns to scale, though we fail to reject operation at constant returns for any individual Fed offices.

The next section briefly discusses the findings of previous studies of the efficiency of Fed check-processing since the implementation of pricing. Section 3 presents our statistical model. Section 4 describes Federal Reserve check processing and our data. Section 5 presents results. Section 6 concludes.

2. Previous studies

The first studies of the efficiency of Federal Reserve check processing after the implementation of pricing mandated by the Monetary Control Act concluded that the pricing regime had improved resource allocation in the processing of checks. For example, whereas Humphrey (1981) found evidence of scale diseconomies at large Fed check offices during the 1970s, Humphrey (1985) found that by 1983 no Fed office experienced diseconomies, and concluded that "the pricing of the Federal Reserve's check service has clearly improved resource allocation for society as a whole" (p. 49).

More recent studies of the Fed's efficiency have tended to support Humphrey's (1985) findings about scale efficiency, but nevertheless conclude that Federal Reserve check operations suffer from considerable cost, or "x-", inefficiency. Using quarterly data for 1979–1990, and both parametric and non-parametric methods, Bauer and Hancock (1993) found evidence of considerable cost inefficiency at Fed check offices both before and after the implementation of pricing, with no significant difference in average inefficiency between the two periods. Further, they concluded that during 1983–1990, Fed check facilities experienced a slight, though statistically insignificant, decline in average productivity.

Bauer and Ferrier (1996) used quarterly data for 1990–1994 to estimate cost functions for Fed check processing, wire transfer, and ACH transfer services. For check processing, Bauer and Ferrier (1996) found that both average cost inefficiency and the dispersion of inefficiency across Fed offices were high during this period. Further,

[2] See Berger and Humphrey (1997) for a recent survey of efficiency studies for banks and other financial institutions.

Bauer and Ferrier (1996) detected evidence of technological regress in check processing during the early 1990s, which they associated with declining processing volume at some sites, migration of "high-quality" check business (e.g., social security and payroll checks) to ACH, and the implementation of new check services (e.g., application of magnetic ink character recognition).

The evidence presented by Bauer and Hancock (1993) and by Bauer and Ferrier (1996) suggests that the efficiency with which the Fed provides check processing services did not improve with the implementation of the pricing regime. The Fed has retained significant market share in the processing of checks, however, and its volumes continued to rise through 1999. Further, although both studies employed fairly flexible methods to estimate cost efficiency, and report results that are robust to different methods, alternative estimation methods exist that offer even more flexibility. Hence, the Fed's continued presence in check clearing, the benefit of additional years of data since the advent of pricing to study the efficiency with which the Fed provides payments services, and the availability of flexible estimation methods (and newly developed means of testing hypotheses based on those methods) justify and enable a new look at the productivity of the Fed's provision of check clearing services. [3]

3. Model and estimation method

We use data envelopment analysis (DEA), which is a non-parametric distance function estimator, to estimate the productivity of Federal Reserve offices in providing check clearing services. DEA has been used widely to study efficiency, but almost never with any attempts at statistical inference. [4] Indeed, DEA and similar estimators are often said to be deterministic or non-stochastic. They are, however, actually *estimators* of unknown distance functions and, consequently, statistical inference is necessary to learn what an *estimate* might reveal about a true distance. Recently, methods of statistical inference have been developed for DEA and similar estimators. In the present context, these methods permit us to discriminate among alternative models of Federal Reserve check production, to test for economies of scale, and to test for differences in productivity across Fed offices.

We begin by defining a production set that gives the set of feasible combinations of inputs and outputs. The boundary of this set is frequently referred to as the *technology* or the *production frontier.* Points on the frontier are regarded as technically efficient, while points in the interior of the production set are technically inefficient; units operating in the interior of the production set could reduce input quantities without reducing output quantities. For points inside the production set, we use the Shephard (1970)

[3] In addition to Bauer and Ferrier (1996), Adams et al. (2002), Bauer and Hancock (1995) and Hancock et al. (1999) have also recently examined efficiency or scale economies in Federal Reserve electronic payments services.

[4] Recent applications of DEA in banking studies include Isik and Hassan (2002), McKillop et al. (2002), Rezvanian and Mehdian (2002) and Sathye (2001). See Berger and Humphrey (1997) for a comprehensive survey of earlier studies.

input distance function to measure distance to the frontier, and estimate this distance using linear programming methods as described in Simar and Wilson (2000b).

We make standard assumptions about the production set to enable estimation of inefficiency: (i) The production set is convex and closed. (ii) All production requires the use of some inputs, and both inputs and outputs are strongly disposable. (iii) The observed set of inputs and outputs for check-processing offices results from independent draws from a probability density function with bounded support over the production set. (iv) This density is strictly positive for all points along the frontier. (v) Starting from any point along the frontier the density is continuous in any direction toward the interior of the production set. Together, these assumptions define the data-generating process that produces the sample observations, and permit statistical estimation and inference about the unobserved technology as well as the unobserved Shephard input distance function.

Unfortunately, few results exist on the sampling distribution of the DEA distance function estimator; in particular, the sampling distribution is known only for the special case of one input and one output. The estimator's distribution remains unknown for cases involving more than one input and one output, making statistical inference by conventional methods impossible in more general cases. [5] However, bootstrap methods described in Simar and Wilson (1998, 2000a,b) allow one to approximate the asymptotic distribution of distance function estimators in multivariate settings, and hence to make inferences about the corresponding *true* distance functions. [6] A second complication arises because non-parametric distance function estimators are biased in finite samples; in particular, our input distance function estimator is biased downward. Intuitively, the bias arises because the estimate of the production frontier is based on actual observations, so in finite samples the location of the estimated frontier will lie on or below the true frontier (assuming no measurement er-

[5] Korostelev et al. (1995) prove that the convex hull of the free disposal hull of the sample observations is a consistent estimator of the production set under conditions met by our assumptions, and Kneip et al. (1998) prove consistency of our distance function estimator under these assumptions and establish its convergence rate of $O_p(n^{-2/(p+q+1)})$, where p denotes the number of inputs and q denotes the number of outputs. The rate of convergence is slow, as is typical in non-parametric estimation; the rate becomes even slower as $p+q$ is increased – this is the well-known curse of dimensionality that commonly plagues non-parametric estimators. The free disposal hull estimator used by Bauer and Hancock (1993) relaxes the convexity assumption, but otherwise is similar to our DEA estimator. Imposing the convexity assumption yields a slightly faster rate of convergence; see Park et al. (2000) for details. Gijbels et al. (1999) derive the asymptotic distribution of the output distance function estimator corresponding to our input-oriented estimator for the special case of one input and one output, along with an analytic expression for its large sample bias and variance; these results easily extend to our input distance function estimator. Unfortunately, derivation of similar results for the more general multivariate setting is complicated by the radial nature of the distance functions and the complexity of the estimated frontier.

[6] The usual, naive bootstrap method where one resamples uniformly, independently, and with replacement from the original sample observations does not lead to consistent inference in the DEA setting due to the bounded sample space. The methods proposed by Simar and Wilson address this problem by resampling instead from a smooth, non-parametric estimate of the density of the sample observations. Kernel density estimators are used, along with a multivariate extension of the reflection method described by Silverman (1986) to avoid bias in the kernel density estimator near the boundary along the estimated frontier.

ror). Hence, estimates of the productivity of individual offices will be biased upward. Fortunately, this bias can be corrected using the heterogeneous bootstrap method of Simar and Wilson (2000a). [7] We report both original and bias-corrected estimates for comparison.

4. Federal reserve check processing

The clearing of checks involves receiving checks from depositing banks (defined broadly to include all depository institutions), sorting them, crediting the accounts of the depositing banks, and delivering the checks to the banks upon which they are drawn. Such "forward item" processing is the main source of revenue and total cost for Fed check operations. Some Fed offices process Federal Government checks and postal money orders, as well as commercial checks. Federal Reserve offices also process "return items" (which include checks returned on account of insufficient funds) and provide various electronic check services, such as imaging and truncation. Finally, Fed check offices entail costs associated with making adjustments necessitated by processing and other errors. Following the convention of other studies, we focus here on the forward processing of commercial and Federal Government check items.

The methods we use permit the estimation of productivity of check offices with multiple outputs. In addition to treating the number of forward items processed as an output, we consider whether the number of endpoints served by an office should be treated as a second output. An endpoint is an office of a depository institution to which a Fed office delivers check items. Other studies have suggested that differences in the number or location of endpoints may help explain why some Fed check offices appear less efficient than others. The number of endpoints (or a measure of the location of endpoints) could be treated as an environmental characteristic affecting the efficiency of check processing. Alternatively, the number of endpoints might be thought of as a measure of the level of service provided by a check office – an office serving many endpoints, all else equal, is providing a higher level of service than an office serving fewer endpoints. In this sense, check processing is analogous to the delivery of mail by a post office. The output of a post office is not simply the number of items it delivers, but also the number of addresses to which it delivers mail. Presumably, a post office that delivers mail to a single address provides less service than a post office that delivers an identical quantity of mail to several addresses.

We test our hypothesis that the number of endpoints served by a Fed check office constitutes a distinct output by first estimating the productivity of Fed check offices for both a single-output model (number of forward check items processed), as well as a two-output model that includes the number of endpoints as a second output. We perform a statistical test to determine whether the data support the treatment of the

[7] Given an initial estimate $\hat{\delta}$ of δ (distance from the frontier) and a corresponding set of bootstrap estimates $\hat{\delta}_b^*$, $b = 1, \ldots, B$, a bias-corrected estimator of δ can be constructed by subtracting the bootstrap bias estimate $(B^{-1}\sum_{b=1}^{B}\hat{\delta}_b^* - \hat{\delta})$ from $\hat{\delta}$. The bootstrap bias estimate is the empirical analog of $E(\hat{\delta}) - \delta$.

Table 1
Definitions and measurement of inputs

1. Personnel – number of employee work hours.
2. Materials, software, equipment and support – expenditures are deflated by the following price measures:
   - *Materials*. GDP implicit price deflator (sa, 1996 = 100);
   - *Software*. Private non-residential fixed investment deflator for software (sa, 1996 = 100);
   - *Equipment*. For 1979–1989, PPI for check-handling machines (June 1985 = 100); for 1990–1999, PPI for the net output of select industries–office machines, n.e.c. (nsa, June 1985 = 100);
   - *Other support*. GDP implicit price deflator (sa, 1996 = 100).
3. Transit – expenditures for shipping, travel, communications, and data communications support deflated by the following price measures:
   - *Shipping and travel*. Private non-residential fixed investment deflator for aircraft (sa, 1996 = 100);
   - *Communications and communications support*. Private non-residential fixed investment deflator for communications equipment (sa, 1996 = 100).
4. Facilities – expenditures on facilities support deflated by the following price index: "Historical Cost Index" from Means Square Foot Costs Data 2000 (R.S. Means Company: Kingston, MA), pp. 436–442. Data are January values.

Sources: Federal Reserve Planning and Control System documents unless otherwise noted. Additional details are available from the authors.

number of endpoints served as a distinct output. Our data consist of quarterly observations for each Federal Reserve Bank main office, branch office, and dedicated check processing center from 1980:Q1 through 1999:Q4, totaling 3761 office-quarters. [8]

Federal Reserve check facilities use a variety of inputs to process checks and deliver them to paying banks. Estimation of productivity using statistical methods requires the specification of a model of the production process with a limited number of inputs. We follow the convention of other studies of check office productivity (Bauer and Hancock, 1993; Bauer and Ferrier, 1996) by defining four distinct categories of inputs used in the processing of forward items: (1) personnel; (2) materials, software, equipment and support; (3) transit services; and (4) facilities. Our model of productivity requires estimates of the physical quantities used of each input, rather than total expenditures. Table 1 describes our method of constructing measures of the four inputs for each Fed check office using expense data for forward items processing. Table 2 gives summary statistics for both inputs and outputs.

5. Empirical results

We estimate input distance functions for both the one- and two-output models described above using pooled cross-section, time-series data on all Fed offices that process checks. [9] Pooling is necessary to obtain meaningful estimates because of the slow convergence rates of our non-parametric estimator. We assume no technical

[8] Quarterly data on the number of items processed and number of endpoints served by each office are from an internal Federal Reserve database containing Federal Reserve expense reports.

[9] The panel is not balanced because two offices closed and one opened during the period covered by our study.

Table 2
Summary statistics for inputs, outputs (3761 observations)

Variable	Mean	SD	Minimum	Maximum
Outputs				
Checks processed (000s)	83 376.11	774.99	10 413.00	280 006.00
Endpoints	412.66	4.67	32.00	1686.00
Inputs				
Personnel	34 553.67	396.25	4905.04	201 529.15
Materials, etc.	1914.33	19.32	154.83	10 630.28
Transit	720.73	9.58	62.76	4438.23
Facilities	1262.45	14.93	20.34	6677.62

regress and, thus, distance function estimates for a particular observation measure distance to the estimated boundary of the production set in the last sample period – 1999:Q4. Consequently, distance function estimates reflect *efficiency*, which relates a unit's performance to the current technology, only for that quarter. Changes in the distance function estimates for an office over time, however, reflect changes in the *productivity* of that office.

Because our estimator of the production frontier is based on the convex hull of the free-disposal hull of the sample observations, the distance function estimates are not independent of one another, and a single outlier has the potential to severely distort estimates for possibly many observations. Using the outlier-detection technique described in Simar (2003), however, we found no evidence of outliers in our data that might serve to distort estimates of the production frontier.

5.1. Tests of specification and returns to scale

First, we examine our conjecture that the number of endpoints served by a check office is a distinct output of check production in addition to the number of items processed. If the number of endpoints is in fact irrelevant, including it would have no influence on the shape of the estimated production frontier. Fig. 1 plots the bias-corrected distance function estimates for 1999:Q4 from the two-output model (model #2) as a function of the corresponding estimates from the single-output model (model #1). The same scale is used on both axes to facilitate comparison. [10] If the esti-

[10] Because the bias-corrected estimates of the distance function are obtained by subtracting a potentially noisy estimate of bias from the original distance function estimates, the bias-corrected estimates might have higher mean-square error than the original estimates. To check this, we computed the value 1/3 times the square of the bootstrap bias estimate divided by the sample variance of the bootstrap estimates, which serves as an indicator of whether mean-square error is worsened when the bootstrap bias estimate is subtracted from the original estimate to obtain the bias-corrected estimate. As discussed in Simar and Wilson (2000a), this ratio should exceed unity if the bias-corrected estimator is to be used; otherwise, the bias-corrected estimator will likely have greater mean-square error than the original, uncorrected estimator. In every case, the ratio is well above unity, and so we rely on the bias-corrected estimator of the distance function.

mates from each model were identical, the points in Fig. 1 would fall on a 45° line running from the lower-left to the upper-right corner of the figure. Several points lie below the 45° diagonal, however, indicating that treating the number of endpoints as an output increases the estimated efficiency of some sites. [11] Treating endpoints as an output also changes the shape of the estimated frontier, which would not be expected if endpoints were irrelevant to the production process.

A formal statistical test using distance function estimates for all quarters of the sample further indicates that the number of endpoints should be treated as a distinct output. The test of the null hypothesis of one output (forward items processing) against the alternative hypothesis that Fed check office production is more appropriately modeled as involving two outputs (with number of endpoints as the second output) is based on the idea that under the null hypothesis, the irrelevant output will be unrelated to the true production frontier. Statistics for the test are based on ratios and differences of distance function estimates from the two models; under the null, the statistics are expected to be small in value, i.e., the distance functions will be similar. To carry out the test, which involves drawing inferences about the true distance function estimates for each observation, we use the heterogeneous bootstrap described in Simar and Wilson (2000a) to approximate the distribution of the true distance function estimator and thereby derive *p*-values for the test. [12] Using 2000 bootstrap replications to obtain *p*-values, our test rejects the one-output model in favor of the two-output model with a *p*-value less than 0.0005 in each case. [13] Hence, the evidence supports the hypothesis that number of endpoints served constitute a distinct output of Federal Reserve check processing.

Next, we investigate returns to scale in Fed check processing. We test the null hypothesis of globally constant returns to scale in the technology of the two-output model versus the alternative hypothesis of variable returns to scale. Using the six statistics described by Simar and Wilson (2002) and 2000 bootstrap replications, we reject the null hypothesis of constant returns with *p*-values of less than 0.001 for each statistic. We are unable to reject the hypothesis of operation at constant returns for any individual Fed office, however, and hence our results conform with conclusions about scale economies in Humphrey (1985) and other studies. Although output

[11] By construction, the uncorrected distance function estimates from model #2 are less than or equal to the corresponding uncorrected distance function estimates from model #1 due to the increased dimensionality in model #2. This is not true for the bias-corrected estimates, however.

[12] We use the heterogeneous bootstrap, rather than the homogeneous bootstrap of Simar and Wilson (1998), because the latter requires the true inefficiency estimates to be statistically independent of output levels and input mix. Using two bootstrap versions of the Kolmogorov–Smirnov test described in Wilson (2003), we reject the null hypothesis of independence for both the one- and two-output models with *p*-values of less than 0.0001.

[13] For each statistic, we find no bootstrap values among the 2000 bootstrap replications that are smaller than the original value computed from the sample; hence the estimated *p*-value is less than $1/2000 = 0.0005$. Details of this test are presented in Simar and Wilson (2001).

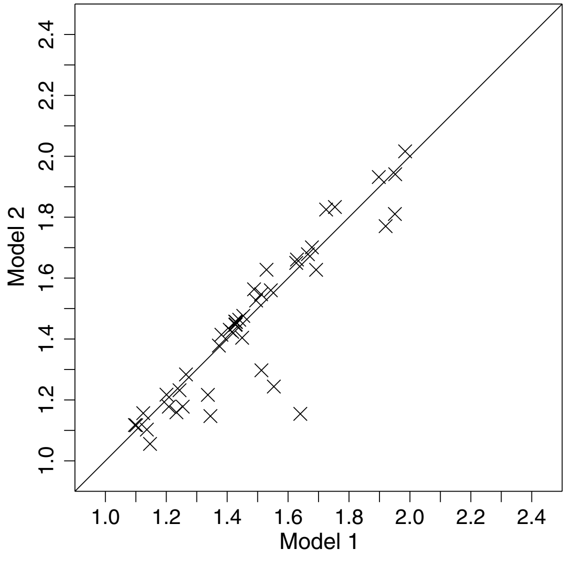

Fig. 1. Model #2 versus model #1 distance function estimates, 1999:Q4.

quantities vary widely among Fed offices, we find no evidence that any office operates at an inefficient scale. [14]

5.2. Productivity change

Changes in distance function estimates for an office over time reflect changes in the productivity of that office. Aggregating across all offices provides evidence on how productivity changed for the system as a whole. Fig. 2 plots the median and variance of the bias-corrected distance function estimates from the two-output model for each quarter of the sample. The median is measured on the left vertical axis, while variance is measured on the right vertical axis. The median varies considerably over the sample period, but after increasing in the mid-1980s, it tends to decline over the remaining years through the 1990s. Because the median distance function estimates are smaller during much of the 1990s than before 1982, our results suggest that the

[14] As with the test for an irrelevant output, our test of returns to scale is based on the idea that under the null hypothesis, distance function estimates obtained while imposing constant returns should not differ greatly from corresponding estimates obtained without imposing constant returns. Although the Monte Carlo experiments in Simar and Wilson (2002) were based on the homogeneous bootstrap to reduce computational burden, our tests here are based on the heterogeneous bootstrap, having rejected independence between the input distance function values and the set of outputs and input angles.

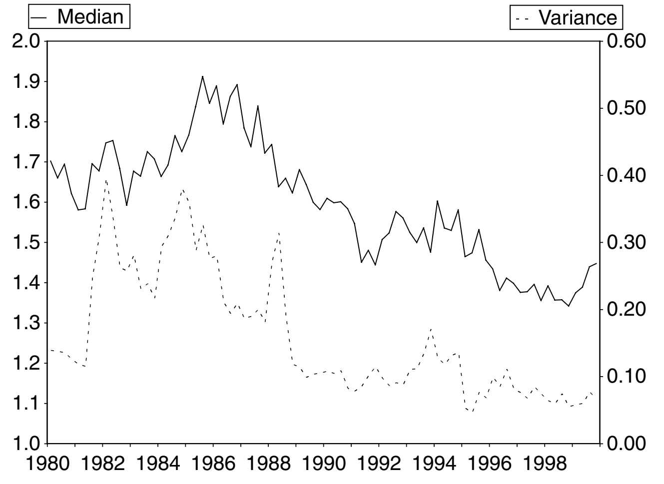

Fig. 2. Median and variance of bias-corrected productivity estimates across Fed check-processing sites, model 2 (two outputs), 1980:Q1–1999:Q4.
Note: Solid line shows median productivity, measured on the left vertical axis; dashed line shows variance of estimated productivity, measured on the right vertical axis.

median productivity of Federal Reserve check offices was higher by the end of our sample than before the implementation of pricing in 1982. We find that median productivity worsened initially after the implementation of pricing, however, consistent with the findings of Bauer and Hancock (1993). Finally, we find that the dispersion in productivity across offices was considerably smaller by the end of the sample period than it had been during the 1980s. It appears that pricing, and perhaps other factors, have narrowed productivity differences across Fed offices. [15]

Table 3 reports input distance function estimates from the two-output model for each Federal Reserve check-processing office in 1999:Q4, the final period of our sample. The column labeled $\hat{\delta}$ gives the original distance function estimates, while the column labeled $\hat{\hat{\delta}}$ gives the bias-corrected distance function estimates obtained as described in footnote 7. The remaining columns of Table 3 contain estimated upper and lower bounds (a^*_α, b^*_α) for confidence intervals at $\alpha = 0.1$ and 0.05 significance levels, respectively, which were obtained using the methods described in Simar and Wilson (2000a). [16]

[15] Even though we reject the one-output model in favor of the two-output model, median distance function estimates based on the one-output model (plotted in Fig. 3) show the same trend as estimates for the two-output model. These results further support our conclusion that the System's median productivity improved beginning in the mid-1980s through the 1990s.

[16] Note that the original estimate of the input distance function, $\hat{\delta}$, lies outside the corresponding estimated confidence interval in each case. As discussed in Simar and Wilson (2000a), the confidence interval estimates incorporate an implicit bias correction that does not depend on an explicit estimate of the bias. The original distance function estimates always lie to the left of the corresponding confidence interval estimates, reflecting the downward bias of the input distance function estimator.

Table 3
Input distance function estimates for model #2, 1999:Q4

Site	$\hat{\delta}$	$\hat{\hat{\delta}}$	$a^*_{0.1}$	$b^*_{0.1}$	$a^*_{0.05}$	$b^*_{0.05}$
1	1.0000	–	–	–	–	–
2	1.0192	1.0556	1.0412	1.0711	1.0377	1.0727
3	1.0587	1.1026	1.0854	1.1216	1.0817	1.1257
4	1.3312	1.3784	1.3586	1.3985	1.3552	1.4026
5	1.5137	1.5600	1.5402	1.5844	1.5361	1.5878
6	1.1668	1.2318	1.2071	1.2556	1.2034	1.2597
7	1.1585	1.2166	1.1925	1.2402	1.1895	1.2464
8	1.0000	1.1168	1.0887	1.1429	1.0826	1.1474
9	1.1113	1.1779	1.1498	1.2098	1.1440	1.2168
10	1.0631	1.1470	1.1147	1.1802	1.1089	1.1844
11	1.0000	1.1176	1.0852	1.1504	1.0770	1.1554
12	1.7265	1.8105	1.7775	1.8429	1.7721	1.8497
13	1.3568	1.4482	1.4131	1.4814	1.4073	1.4875
14	1.3552	1.4463	1.4131	1.4816	1.4076	1.4881
15	1.6024	1.7010	1.6666	1.7357	1.6608	1.7413
16	1.1458	1.2173	1.1834	1.2523	1.1787	1.2595
17	1.0591	1.1565	1.1204	1.1903	1.1131	1.1955
18	1.8110	1.9319	1.8940	1.9700	1.8863	1.9754
19	1.2240	1.2973	1.2622	1.3381	1.2558	1.3464
20	1.3158	1.4041	1.3633	1.4425	1.3544	1.4511
21	1.2036	1.2837	1.2472	1.3266	1.2366	1.3310
22	1.3113	1.4303	1.3898	1.4720	1.3772	1.4768
23	1.0455	1.1593	1.1169	1.2004	1.1085	1.2098
24	1.5487	1.6610	1.6190	1.7031	1.6094	1.7096
25	1.4242	1.5265	1.4833	1.5714	1.4767	1.5785
26	1.3553	1.4558	1.4118	1.5001	1.4067	1.5076
27	1.5254	1.6499	1.6028	1.6917	1.5905	1.7029
28	1.0913	1.1779	1.1345	1.2244	1.1277	1.2322
29	1.5590	1.6784	1.6322	1.7225	1.6230	1.7302
30	1.0477	1.1542	1.1097	1.2034	1.1020	1.2157
31	1.3593	1.4573	1.4069	1.5085	1.3993	1.5173
32	1.2679	1.4140	1.3619	1.4626	1.3522	1.4731
33	1.8130	1.9411	1.8890	1.9914	1.8778	2.0005
34	1.1202	1.2439	1.1919	1.2972	1.1844	1.3079
35	1.2794	1.4236	1.3673	1.4826	1.3578	1.4942
36	1.3489	1.5632	1.4907	1.6260	1.4710	1.6367
37	1.3232	1.5469	1.4724	1.6126	1.4604	1.6237
38	1.0000	1.8251	1.7375	1.8780	1.7099	1.8819
39	1.3494	1.4760	1.4023	1.5546	1.3912	1.5662
40	1.6042	1.8335	1.7466	1.9168	1.7342	1.9334
41	1.2378	1.4633	1.3776	1.5469	1.3595	1.5584
42	1.0000	1.6274	1.5299	1.7056	1.5083	1.7160
43	1.6600	2.0162	1.8796	2.1477	1.8508	2.1678
44	1.4103	1.7706	1.6127	1.9443	1.5812	1.9844
45	1.2245	1.6272	1.4316	1.8607	1.4050	1.9212

The bias-corrected distance function estimates shown in Table 3 range from 1.0556 to 2.0162, with a mean of 1.4528. Hence, our estimates indicate that in 1999:Q4 the average Federal Reserve check processing site could have feasibly

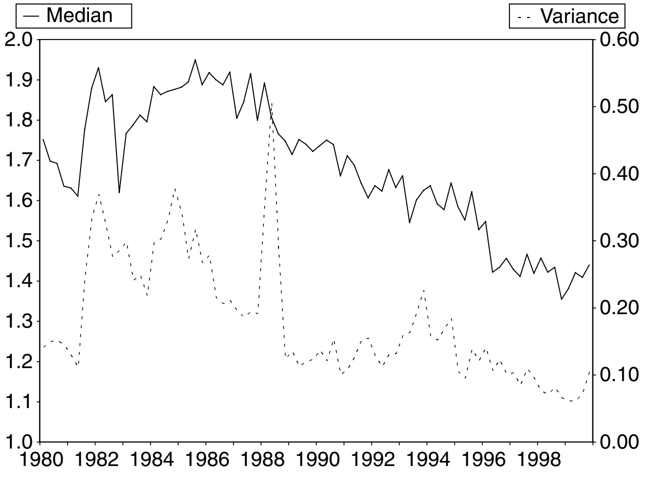

Fig. 3. Median and variance of bias-corrected productivity estimates across Fed check-processing sites, model 1 (one output), 1980:Q1–1999:Q4.
Note: Solid line shows median productivity, measured on the left vertical axis; dashed line shows variance of estimated productivity, measured on the right vertical axis.

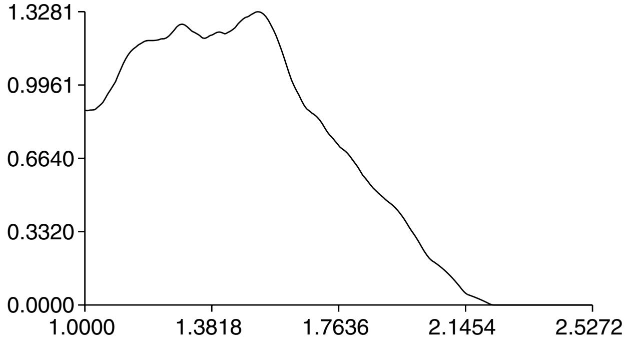

Fig. 4. Kernel estimate of density of $\hat{\hat{\delta}}$ 1999:Q4.

reduced its inputs proportionately by $(1 - \frac{1}{1.4528}) \times 100 = 31.17\%$ without reducing output. [17] As is typical in efficiency estimation, however, the distribution of efficiency estimates is skewed. Fig. 4 shows a non-parametric kernel estimate of the density of

[17] Our estimates of inefficiency are in line with what others have found for commercial banks and Fed services (see Berger and Humphrey, 1997). By contrast, when using a free disposal hull estimator, Bauer and Hancock (1993) estimate average inefficiency of only about 2%. Bauer and Hancock estimate separate frontiers for each quarter and each year, however, and their finding that well over half of all check offices lie on the estimated frontier indicates that there are too few cross-sectional observations to produce meaningful non-parametric estimates of the efficient frontier. We avoid this "curse of dimensionality" by pooling the data in our sample, and obtain average inefficiency estimates that are similar to what Bauer and Hancock (1993) and Bauer and Ferrier (1996) found using parametric techniques.

the bias-corrected distance function estimates for 1999:Q4. [18] In addition to skewness, the estimated density first increases as one moves to the right from 1.0, before eventually decreasing on the right. Hence, in terms of technical efficiency, processing sites are not clustered along the frontier; rather, only a few sites define the frontier estimate, with most lying in the interior of the estimated production set. Our results indicate, therefore, that despite improvement over time, at the end of the sample period many Fed check offices remained considerably less efficient than they could be.

6. Summary and conclusions

The Monetary Control Act of 1980 sought, among other things, to improve the efficiency with which the Federal Reserve provides payments services by requiring Fed offices to recover their costs of providing services plus a private sector adjustment factor. Although prior studies found that Fed offices operated at efficient scale in processing checks after the introduction of the pricing requirement, they also concluded that the introduction of pricing produced no improvement in overall operating efficiency. When the Fed reconsidered its role in the payments system in 1998, one objective it stated for continuing to provide retail payments services was to improve the efficiency of the payments system. Hence, evidence that Fed offices waste resources in the processing of checks would indicate that the Fed could contribute to the efficiency of the payments system by improving the efficiency of its own operations.

The present study reports new evidence on the productivity of Federal Reserve check offices based on non-parametric estimation and recently developed methods of statistical inference for non-parametric estimators. Like prior studies, we treat forward check items processing as an output of Fed offices. However, we also treat the number of endpoints served by an office as a distinct, second output. Our specification tests indicate that treating the number of endpoints served as an output is appropriate. The more endpoints a check office serves, all else equal, the higher the level of service it provides. Failure to treat the number of endpoints served as an output could lead to biased estimates of efficiency, specifically making offices that serve high numbers of endpoints appear less efficient than those serving few endpoints.

We find that median productivity of Fed check offices has improved markedly since the implementation of pricing in the early 1980s, though most of the improvement has come since the late 1980s, after some regress in the middle part of that decade. Further, the variance in productivity across offices has also declined substantially.

[18] The bandwidth for the density estimate was selected using a modified version of the normal reference rule suggested by Hjort and Jones (1996), which incorporates information from the third and fourth sample moments of the data; the bandwidth used was 0.2078. The density estimate was constructed using the reflection method described by Silverman (1986) to overcome the problem of bias near the left boundary.

Unlike previous studies, we report robust confidence intervals around inefficiency estimates for individual Fed offices. Hence, we are able to test hypotheses about differences in inefficiency across offices. We find that although the variance across offices has declined over time while median productivity has improved, significant differences across offices remain. We find that few offices operate close to the efficient frontier, suggesting that further improvements in efficiency are possible at many offices. Moreover, we find evidence of scale economies in check processing, though we cannot reject operation at constant returns for any individual office. Thus, the Fed's recent decision to reduce its check processing capacity by eliminating check operations at some offices appears consistent with the efficiency objectives of the Monetary Control Act, though a definitive conclusion must await an ex post analysis. Such an analysis could be especially useful to policymakers if check volume continues to decline, as one might expect from the growing popularity of electronic payments media and the likely enactment of legislation facilitating the use of electronic check images for the final settlement of payments, and the Fed is forced to consider additional capacity reductions to ensure that its check operations remain competitive.

References

Adams, R.M., Bauer, P.W., Sickles, R.C., 2002. Scale Economies, Scope Economies, and Technical Change in Federal Reserve Payments Processing. Board of Governors of the Federal Reserve System, Washington, DC.

Bauer, P.W., Ferrier, G.D., 1996. Scale economies, cost efficiencies, and technological change in Federal Reserve payments processing. Journal of Money, Credit, and Banking 28, 1004–1039.

Bauer, P.W., Hancock, D., 1995. Scale economies and technological change in Federal Reserve ACH payment processing. Federal Reserve Bank of Cleveland Economic Review (Quarter 3) 14–29.

Bauer, P.W., Hancock, D., 1993. The efficiency of the Federal Reserve in providing check processing services. Journal of Banking and Finance 17, 287–311.

Berger, A.N., Humphrey, D.B., 1997. Efficiency of financial institutions: International survey and directions for future research. European Journal of Operational Research 98, 175–212.

Board of Governors of the Federal Reserve System, 2001. Annual Report, Washington, DC.

Federal Reserve System, 2002. Retail Payments Research Project. Washington, DC.

Federal Reserve System, 2003. Financial Services Policy Committee, press release, February 6, 2003. Available from <www.federalreserve.gov/boarddocs/press>.

Gijbels, I., Mammen, E., Park, B.U., Simar, L., 1999. On estimation of monotone and concave frontier functions. Journal of the American Statistical Association 94, 220–228.

Hancock, D., Humphrey, D.B., Wilcox, J.A., 1999. Cost reductions in electronic payments: The roles of consolidation, economies of scale, and technical change. Journal of Banking and Finance 23, 391–421.

Hjort, N.L., Jones, M.C., 1996. Better rules of thumb for choosing bandwidth in density estimation. Department of Mathematics, University of Oslo, Oslo, Norway.

Humphrey, D.B., 1981. Economies to scale in Federal Reserve check processing operations. Journal of Econometrics 15, 155–173.

Humphrey, D.B., 1985. Resource use in Federal Reserve check and ACH operations after pricing. Journal of Bank Research 16, 45–53.

Isik, I., Hassan, M., 2002. Technical, scale and allocative efficiencies of Turkish banking industry. Journal of Banking and Finance 26, 719–766.

Kneip, A., Park, B.U., Simar, L., 1998. A note on the convergence of nonparametric DEA estimators for production efficiency scores. Econometric Theory 14, 783–793.

Korostelev, A., Simar, L., Tsybakov, A., 1995. On Estimation of Monotone and Convex Boundaries. Publications de l'Institut de Statistique de l'Université de Paris XXXIX 1, pp. 3–18.

McKillop, D.G., Glass, J.C., Ferguson, C., 2002. Investigating the cost performance of UK credit unions using radial and non-radial efficiency measures. Journal of Banking and Finance 26, 1563–1591.

Park, B., Simar, L., Weiner, C., 2000. The FDH estimator for productivity efficiency scores: Asymptotic properties. Econometric Theory 16, 855–877.

Rezvanian, R., Mehdian, S., 2002. An examination of cost structure and production performance of commercial banks in Singapore. Journal of Banking and Finance 26, 79–98.

Rivlin, A.M., Kelley, E.W., McDonough, W.J., Melzer, T.C., 1998. The Federal Reserve in the Payments Mechanism. Committee on the Federal Reserve in the Payments Mechanism, Federal Reserve System, Washington, DC.

Sathye, M., 2001. X-efficiency in Australian banking: An empirical investigation. Journal of Banking and Finance 25, 613–630.

Shephard, R.W., 1970. Theory of Cost and Production Functions. Princeton University Press, Princeton, NJ.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.

Simar, L., 2003. Detecting outliers in frontier models: A simple approach. Journal of Productivity Analysis, in press.

Simar, L., Wilson, P.W., 1998. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. Management Science 44, 49–61.

Simar, L., Wilson, P.W., 2000a. A general methodology for bootstrapping in non-parametric frontier models. Journal of Applied Statistics 27, 779–802.

Simar, L., Wilson, P.W., 2000b. Statistical inference in nonparametric frontier models: The state of the art. Journal of Productivity Analysis 13, 49–78.

Simar, L., Wilson, P.W., 2001. Testing restrictions in non-parametric efficiency models. Communications in Statistics 30, 159–184.

Simar, L., Wilson, P.W., 2002. Nonparametric tests of returns to scale. European Journal of Operational Research 139, 115–132.

Wilson, P.W., 2003. Testing Independence in Models of Productive Efficiency. Department of Economics, University of Texas, Austin, TX.