Competition and Regulation in the Postal and Delivery Sector

Edited by

Michael A. Crew

Center for Research in Regulated Industries (CRRI), Rutgers, The State University of New Jersey, Newark, USA

Paul R. Kleindorfer

The Wharton School, University of Pennsylvania, USA and INSEAD, France

ADVANCES IN REGULATORY ECONOMICS

Edward Elgar

Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
Glensanda House
Montpellier Parade
Cheltenham
Glos GL50 1UA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

ISBN 978 1 84720 507 0

Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

21. Are there economies of scale in mail processing? Getting the answers from a large-but-dirty sample*

Lawrence Fenster, Diane Monaco, Edward S. Pearsall and Spyros Xenakis

1. INTRODUCTION

In this chapter we present econometric evidence that United States Postal Service (USPS) mail processing plants are mostly operated at levels where the returns to density and scale are decreasing. The evidence is derived from production functions fit as stochastic switching regressions to large panel samples of pieces, work hours, capital usage, delivery points, delivery units and other plant-level information mostly drawn from USPS's Management Operating Data System (MODS). Decreasing returns were found for production functions defined for every aggregate of pieces handled by shape, for pieces fed in most single automated, mechanical and manual processes, and when piece-handlings were divided into inbound and outbound sub-streams.

Samples drawn from MODS are problematic because they exhibit anomalies at frequencies suggesting that they are a dirty mix of good observations and occasional reporting mistakes. Our econometrics employs a Maximum Likelihood (ML) estimator for fitting a two-regime stochastic switching regression model as first proposed by Quandt (1972). The motivation for this approach is that the good observations are generated by the production function, while the bad observations are the result of data collection failures consistent with a different regime.

Our estimates indicate that mail processing is primarily an industrial process rather than a network support activity. Returns to scale typically exceed returns to density. This result is anomalous for an activity like transportation that directly supports a distribution network, but it is readily explained for an industrial process as the effect of production smoothing. USPS plants serving larger and more diverse segments of the network apparently benefit from longer continuous flows of mail through the plants.

Our results also show that scale is only one of many statistically significant determinants of productivity among USPS processing plants. Other determinants include the types of delivery points served by a plant, the composition of the mail stream passing through the plant, the ratio of originating to destinating mail in a plant's service area, the configuration of related equipment/processes and the skill profile of a plant's work force.

2. MAIL PROCESSING

USPS plants process mail in separate outgoing and incoming streams for letters and cards, flats, parcels and Priority Mail.[1] Outgoing mail is cancelled, barcoded and sorted for transportation to other processing centers or for insertion into the local incoming stream. Incoming mail is sorted at least to the carrier-route level. Most incoming letter mail is further sorted to the delivery sequence on the routes. Workshared mail is inserted at various points in the processing streams. Processing of letters and flats, and some processing of parcels and Priority Mail, is now largely done with automated and/or mechanized equipment that is specialized by shape. Manual processing is reserved for pieces that either cannot be fed to automated equipment or have been rejected by the automated processes.

Mail processing is typically viewed as a network support activity; first, because sorting the mail is essential for moving it along the USPS network and, second, because processing plants are located at hubs of the network and serve as transshipment points. In testimony before the US Postal Regulatory Commission (PRC), the principal USPS witness compared mail processing to transportation along the network (Bozzo, 2006).

However, there is an alternative way to view mail processing that describes the activity as a conventional industrial process. In this alternative view, mail processing plants are like generating plants on an electrical network. A generating plant is certainly affected by activity on the network, but its output is solely measured by the power it produces. The number of customers or the geographic area served by the electrical network are not considered outputs of the plant. A similar view of mail processing would confine the measures of a plant's output to the volumes and shapes of mail processed. The levels of network variables, such as the number of delivery points and the number of delivery units served by the plant, may have important effects on a plant's productivity, but are not included among the plant's outputs. This view conforms well to the cost model used by USPS and the PRC in postal rate proceedings. The cost model treats mail processing costs as separable by shape with a single driver – piece handlings.

The distinction between mail processing as a network activity and as an industrial process is important because the two views lead to divergent predictions of the effects of expanding the network on the productivity of a plant.

When treated as additions to output, more delivery points and delivery units would lead to a decrease in the volume of mail that is processed by a fixed work force and complement of equipment at a plant. This happens when additional runs are needed to sort to more divisions. Then, more piece handlings are required to sort a given volume of mail. Expanding the network may also disproportionately increase the setup and teardown times for processing runs, especially with automated equipment. For example, more bins would have to be set up and cleared at the beginning and end of each run to sort the same volume of mail.

Both of these effects are likely to be small. The processing schemes used by USPS plants all involve fixed integer numbers of passes through the machines. For example, delivery sequence sorting of letters is mostly done in two passes through the Distribution Bar Code Sorting (DBCS) machines. New delivery points or delivery units are usually accommodated by increasing the number of utilized bins rather than by increasing the number of passes and piece handlings. Adding bins increases setup and teardown time.

But this time is only a small fraction, typically less than 10 percent, of the total time for a processing run.

When mail processing is viewed as an industrial process whose output cannot be stored, the network output variables indirectly represent the limitations on continuous operations imposed by a plant's schedule of arrivals and dispatches. The productivity of electric generating plants improves when the plants are connected to larger networks by the expansion of power grids. Productivity improves because the plants can generate electricity at efficient levels over longer continuous periods of time. Mail processing is also done most efficiently in plants where the various processes can be conducted almost continuously with few interruptions. When production is interrupted, the inevitable result is a decline in productivity due to the temporary forced idleness of both labor and capital.

USPS plants serving larger and more diverse segments of the network are often the beneficiaries of smoother flows of mail through the plant. Mail processing can be more efficient in these plants because the various processes can be conducted in longer uninterrupted runs. A plant serving a small number of delivery points and delivery units will find it hard to schedule arrivals and dispatches to create a continuous flow of pieces through the plant. As the number of delivery points and delivery units grows, it becomes possible (up to a point) to smooth the schedule so that the same volumes pass through the plant over longer uninterrupted periods of time. This can be done by deferring the processing of incoming mail to delivery units that are nearby and to routes, such as suburban residential routes, where deliveries are made later in the day.

Mail processing has aspects of both a network support activity and an industrial process, so the controlling aspect must be ascertained empirically.

3. RETURNS TO DENSITY AND SCALE

For distribution networks it is customary to distinguish between returns to density, which describe the effects of changing inputs proportionately with the network fixed, and returns to scale, which describe the effects with the network variable. Returns to density and scale as defined for the USPS network directly relate to current issues of costing and plant consolidation.

Returns to density and scale in mail processing have been investigated for more than 30 years, with conflicting results. In specific density/scale studies and in studies that have attempted to develop cost-based rates in USPS postal rate proceedings before the PRC, both increasing and decreasing returns have been found.[2]

Beginning in 1997, USPS has presented econometric studies in five successive rate proceedings to support the contention that mail processing labor variabilities are less than 100 percent (US Postal Rate Commission, 1998, 2001, 2002, 2006; US Postal Regulatory Commission, 2007). This evidence has been rejected by the PRC, which continues to use variabilities close to 100 percent based on an assumption that work hours vary mostly in proportion to volumes processed.

Labor variability is defined as the derived demand elasticity of labor with respect to pieces fed or handled in various processing operations. The importance of the variability estimates for US rate-setting is that they are used to derive the marginal costs of processing, which form part of the per unit attributable costs of the various subclasses of mail.

These unit attributable costs have been legislated as rate floors for most subclasses. Marginal costs are also essential components of the Ramsey pricing formulas for welfare-maximizing rates. However, the PRC does not rely upon the formulas for recommending rates.

Variabilities are not the same thing as returns to density or scale, but the concepts are closely related. Variabilities of less than 100 percent are likely to accompany increasing returns to density; decreasing returns to density signal that variabilities exceed 100 percent.

In 2006, USPS sought PRC approval of the models and methods of analysis used in ongoing studies of possible restructurings of its network (US Postal Rate Commission, 2006). The cost functions for these models describe marginal costs for mail processing that decline in steps to approximate the same less-than-100-percent variabilities rejected by the PRC in rate proceedings. The importance of the variability estimates for the USPS network studies is that the declining marginal cost functions will drive the results towards restructurings that consolidate processing plants. This occurs because the cost of processing a given volume of mail, as predicted by the USPS's cost functions, always decreases when consolidation leaves the network with fewer but larger plants.

4. THE MODS DATA

We derive process- and shape-level economies of density and scale from restricted translog production functions fit to panel samples aggregated from MODS.[3] The samples consist of quarterly observations from GFY 1999 to GFY 2005 of man-hours (HRS), total pieces fed (TPF), total pieces handled (TPH) and first-handled pieces (FHP) for various manual and machine-specific operations by shape (letters, flats and parcels) at up to 368 USPS processing plants. The MODS data has been supplemented by capital indices (QICAP) derived from USPS equipment registries and facility records, and by other plant statistics drawn from USPS records by Bozzo, including the numbers of various kinds of delivery points (DP) and delivery units (DU) served by the plant. The MODS data have also been matched to volume statistics from other USPS sources for the areas served. Altogether, the data set includes several alternative measures of processing outputs, matching observations for work hours, capital services and network coverage, and enough information to fashion controls for mail flow composition, plant and process configuration, labor force composition and the distribution of delivery points by type to support econometric modeling of production functions by shape and by process.

In releasing the MODS data, USPS has followed the practice of not providing any information that could be used to specifically identify individual plants. This makes it impossible to supplement the MODS observations with information that could be used to fashion a complete and unambiguous set of controls for exogenous factors affecting processing productivities. Among the factors for which we have no direct controls are the quality of the local postal labor force, the geographic dimensions and demographic characteristics of a plant's service area, the proximity of the delivery units served by the plant, the condition of the local ground transportation system, the plant's air and ground connections, the smoothness (or lack of smoothness) of the plant's arrival and dispatch schedules, the age of the plant, its classification by type, its square foot area, number of floors and other salient physical characteristics. These are major omissions. Nevertheless, the basic MODS observations of piece handlings and work hours have been supplemented by Bozzo with enough plant-level information to allow us to define variables that indirectly control for most of the major factors that exogenously affect productivity.

The MODS data includes work hours and labor costs for several broad Labor Distribution Code (LDC) categories from which nominal labor rates may be derived. However, these nominal rates cannot be converted to real rates, because a local wage or cost-of-living index cannot be associated with the plants without identifying them. The MODS data also contains no information whatsoever that might be used to derive real rental rates for the different categories of equipment and other capital used in the plants. These omissions are not important for fitting production functions, but are serious handicaps when the MODS data are used to fit cost functions or derived demand functions.[4]

Panel samples drawn from MODS are large but known to be dirty because of the methods used to collect the data. MODS is a real-time USPS reporting system with weak quality controls at the point of collection and no mechanism for identifying and recovering missing observations. The missing observations are indistinguishable from observations that are truly zero-valued. The MODS data are collected by work shift and aggregated in stages by week, accounting period and postal quarter, with missing observations treated as zero values at each stage. HRS is subject to various kinds of clocking errors made by employees that mis-assign work hours to processes. QICAP is derived from USPS equipment inventory registries. The major difficulty with these records is that equipment changes are sometimes registered at times that do not correspond exactly to when the equipment actually entered or left service. The other explanatory variables in the data set appear to be much less likely to contain major errors than HRS and QICAP. However, it is still possible that errors in the other variables may cause some of the data to be bad.

Large-but-dirty samples are an increasingly common result of self-collection schemes using electronic equipment. The amount of data that can be collected cheaply by these schemes is often enormous. MODS is typical. The samples we have extracted often contain over 8000 observations. Unfortunately, the practice that makes self-collection cheap is an absence of review and verification as the data are collected. Inevitably, the samples produced by such schemes contain bad observations because self-collection is prone to occasional undetected malfunctions.[5]

The observations of HRS, TPF (or TPH) and FHP for the various processes all contain apparent errors and anomalies that have been described and documented in recent PRC Decisions (US Postal Rate Commission, 1998, 2001, 2002, 2006). The anomalies mostly consist of unrealistic extreme values for TPF/HRS, occasional incompatible values for TPF, TPH and FHP, and missing values for either HRS or TPF (or TPH) but not both. The anomalies are more frequent for processes that are starting up or shutting down at a plant. The patterns suggest that the bad data in the samples are the result of intermittent reporting mistakes.

Bozzo (2006) and Roberts (2006) have attempted to deal with the dirty data by using screens to delete the grossly erroneous observations and, in some cases, by applying an Instrumental Variables (IV) estimator. Both approaches require auxiliary assumptions and non-sample information. Bozzo's screens remove observations from the samples used to fit his equations at frequencies that range from about 0.7 percent to over 20.0 percent. The screens cannot identify and remove all of the bad data in the samples without also non-randomly deleting good data.[6] This risks creating selection biases in the screened sample. Nevertheless, the screens (but not the IV estimator) are compatible with an assumption that the bad data are intermittent. We have not screened the MODS data for our work except to remove observations when the information needed to construct the variables of the production function is incomplete.

5. THE STOCHASTIC SWITCHING REGRESSION MODEL

Our research employs a Maximum Likelihood (ML) estimator to fit two-regime linear stochastic switching regressions with exogenous fixed-probability independent switching.[7] The ML estimator relies only on sample information. The estimates are consistent, asymptotically normal with a known variance–covariance matrix, and asymptotically efficient under general regularity conditions as described in standard statistics and econometrics texts (Wilkes, 1962; Maddala, 1977; Judge et al., 1985).

Linear stochastic switching regressions are a natural choice for modeling large-but-dirty samples that contain an unknown mix of 'bad' and 'good' observations. Bad observations are defined as observations that are not generated by the same linear equation with the same standard error as the more numerous good observations. Ordinary Least Squares (OLS) estimates of the production functions will be inconsistent when the good and bad equations are different if any bad data remains after screening or if the screens non-randomly remove any good data. However, the ML estimator for the stochastic switching model makes it unnecessary to identify and delete bad observations from the sample. Instead, the ML estimates allow us to compute a set of conditional probabilities, one for each observation in the sample, that correspond to the Bayesian posterior probabilities that the observations are bad (or good).

Many of the kinds of problems that are known to cause OLS estimates to be inconsistent may be viewed as the result of samples generated by an observationally equivalent linear model. In Pearsall (2007) it is shown that outliers, replaced observations and errors-in-variables can all be treated as dirty data from an observationally equivalent linear model. The essential aspect of the bad observations for the ML estimator is that they are intermittent in the sample.

In our applications the good observations are assumed to be generated by the translog production function whose parameters and standard error correctly describe the normal operation of USPS processing plants. The bad observations are assumed to be generated by a different linear equation with the same form and variables.

The ML estimator yields estimates for the coefficients and standard errors of the equations for both the bad and good observations, and an estimate of the frequency of the bad data. The estimates were computed using a special-purpose algorithm embedded as a macro in a set of Lotus 1-2-3 worksheets.[8] The algorithm uses the first-order conditions for a maximum of the log-likelihood function within an iterative process to simplify and speed the calculation of the estimates. The method for computing the asymptotic variance–covariance matrix is described in Judge et al. (1985) and Maddala (1976). The conditional probabilities have been computed and are used to estimate sample means for the good observations in the MODS samples. The sample means and the parameter estimates from the good equations are used together to calculate elasticities.

6. THE PRODUCTION FUNCTIONS

Shape-level production functions were fit for letter, flat and parcel sorting operations and for letter cancellation operations. Process-level production functions were fit for automated, mechanized and manual sorting and cancellation operations. Separate production functions were also fit for the inbound and outbound components of the letter and flats sorting operations both by shape and by process.

The processes are identified by the equipment they use:

1. Letter Sorting: Mail Processing Bar Code Sorters (MPBCS), Distribution Bar Code Sorters (DBCS), Optical Character Readers (OCR) and Manual Letter Sorting.
2. Flats Sorting: Flats Sorting Machine 881 (FSM 881), Flats Sorting Machine 1000 (FSM 1000), Automated Flats Sorting Machine (AFSM 100) and Manual Flats Sorting.
3. Cancellation: Advanced Facer Canceller Machine (AFCS) and Non-AFCS Cancellation.
4. Parcels Sorting: Small Parcel Bar Code Sorter Other that Priority Mail (SPBS Other), Small Parcel Bar Code Sorter Priority Mail (SPBS Priority), Manual Parcels Other than Priority Mail and Priority Mail.

The same general equation form was used for the production functions for the shape-level and process-level models. The fitted equations are translogs with respect to the labor and capital inputs and measures of the network served by the plant. The translog was chosen because it is a general form that imposes few structural restrictions on the shape of the fitted production functions. In addition, the equations include sets of other controls to account for exogenous factors affecting the productivity of the plants. These factors are the season of the year, a productivity trend, the distribution of delivery points by type, the composition of the mail stream for the applicable shape, the ratio of originating to destinating mail volume for the shape, the availability of equipment and processes at the plant, and the skill profile of the plant's labor force.

The dependent variables of the regression equations are natural logarithms of pieces fed or handled as follows:

ln(FHP) – the aggregate of first-handled pieces for the shape-level sorting operations;
ln(TPH) – total pieces handled for shape-level cancellations;
ln(TPF) – total pieces fed for automated and mechanized sorting processes; and
ln(TPH) – for manual sorting and cancellation processes.

The MODS data are aggregated to postal quarters of unequal length so weekly averages of the pieces handled (FHP, TPF and TPH), work hours (HRS), and the capital services indices (QICAP) were calculated and used in the regressions. The data for work hours in MODS exactly match the data for piece handlings, but the capital indices constructed by Bozzo match only approximately.[9] Two measures of the size of the distribution network assigned to the plants are used in the production functions. These are the total number of delivery points (DP) in the plant's service area, and the total number of delivery units (DU) assigned to the plant.

The translog component of the production function is as follows:

$$\begin{aligned}
&\alpha + \beta_l \ln(\mathrm{HRS}) + \beta_k \ln(\mathrm{QICAP}) + \beta_p \ln(\mathrm{DP}) + \beta_u \ln(\mathrm{DU}) \\
&\quad + \ln(\mathrm{HRS}) [\beta_{ll} \ln(\mathrm{HRS}) + \beta_{lk} \ln(\mathrm{QICAP}) + \beta_{lp} \ln(\mathrm{DP}) + \beta_{lu} \ln(\mathrm{DU})] \\
&\quad + \ln(\mathrm{QICAP}) [\beta_{kk} \ln(\mathrm{QICAP}) + \beta_{kp} \ln(\mathrm{DP}) + \beta_{ku} \ln(\mathrm{DU})] \\
&\quad + \ln(\mathrm{DP}) [\beta_{pp} \ln(\mathrm{DP}) + \beta_{pu} \ln(\mathrm{DU})] + \beta_{uu} \ln(\mathrm{DU}) \ln(\mathrm{DU}).
\end{aligned}$$

We have defined sets of additional control variables for most of the information that has been matched to the MODS data. The added controls are as follows:

1. *Dummy variables for postal quarters 1, 2 and 3*. These controls represent systematic seasonal variations in productivities due to changes in address quality, work force composition, arrival/dispatch schedules, and so on.
2. *An annual trend*. The mid-quarter time in years from the beginning of the first quarter of GFY 1999. This variable captures the effects of gradual progressive changes in productivity associated with the adaptation of improved technology and changes in the education or training of the work force.
3. *Delivery point shares.* The fraction of DP consisting of various types of delivery points: CURB – city kerb delivery points, NDCBU – city box units, CENT – city central delivery points, OTHER – other city delivery points, RB – rural boxes, HCT – highway contract delivery points, and POBOX – possible P.O. Box deliveries. All except the share for OTHER are included in every equation.
4. *Mail subclass and worksharing shares*. The shares of originating, destinating or total volume by selected subclasses and worksharing categories for the mail shape (letters, flats or parcels) for the territories served by each plant. Shares of total mail for the shapes by Revenue, Pieces and Weight (RPW) category have been computed from USPS volumes passing through the plants. The shares are included in the production functions except for the share of single-piece first-class letters and flats, and the combined share of first-class and priority parcels to account for the effects of the composition of the mail stream on productivity.
5. *ln(origin/destination)*. The natural logarithm of the ratio of originating to destinating volume for the mail shape. The ratio is computed from volume data rather than from piece handlings, to avoid including the equation error in the calculation of the ratio.
6. *Shape dummy variables*. Shape dummy variables for parcels, Priority Mail and/or cancellation were included in many of the equations. These dummies are set to one in plant/quarters for which TPH for all processes involving these shapes are zero. The shape dummies are included in the production functions as a means of identifying the type of processing plant.
7. *Process dummy variables*. Process dummy variables were defined for all of the automated and mechanized machine types. These dummies are set to one for a process under the following two conditions: (1) the plant handles mail of the associated shape, and (2) TPH is zero for the machine type in the plant/quarter. These variables identify the machine types that are missing or inactive in each plant/quarter. The production functions include those process dummy variables that are appropriate for the shape of mail, but omit the dummy for the process itself. The process dummies represent the equipment configuration at the plant for the shape.

8. *Labor skill shares*. The MODS data has been supplemented with plant-level work hours in five LDC categories: Automated letters sorting, Mechanized flats sorting, Mechanized parcels sorting, Manual sorting and Allied labor. This data was used to compute work force shares. Two of these share variables are included in each of the equations to represent the skill distribution of the plant's work force.

7. ESTIMATES OF RETURNS TO DENSITY AND SCALE

We calculate returns to density and scale by summing output elasticities with respect to HRS, QICAP, DP and DU. For our translog production functions these elasticities are linear functions of ln(HRS), ln(QICAP), ln(DP) and ln(DU) that employ the ML estimates of the coefficients for the good equations. The following formulas express returns to density and scale as elasticities of FHP, TPF or TPH with respect to proportionate changes in the input and network variables:

Returns to density: $E_l + E_k = \partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{HRS}) + \partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{QICAP})$.

Returns to scale:[10] $\dfrac{E_l + E_k}{1 - E_p - E_u} = \dfrac{\partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{HRS}) + \partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{QICAP})}{1 - \partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{DP}) - \partial\ln(\mathrm{FHP})/\partial\ln(\mathrm{DU})}$.

For the process-level equations, TPF or TPH appear in the formulas in place of FHP. The elasticities in the formulas are derived from the coefficients of the translog production functions as follows:

$$
\begin{aligned}
E_l &= \beta_l + 2.0\beta_{ll}\ln(\mathrm{HRS}) + \beta_{lk}\ln(\mathrm{QICAP}) + \beta_{lp}\ln(\mathrm{DP}) + \beta_{lu}\ln(\mathrm{DU}),\\
E_k &= \beta_k + \beta_{lk}\ln(\mathrm{HRS}) + 2.0\,\beta_{kk}\ln(\mathrm{QICAP}) + \beta_{kp}\ln(\mathrm{DP}) + \beta_{ku}\ln(\mathrm{DU}),\\
E_p &= \beta_p + \beta_{lp}\ln(\mathrm{HRS}) + \beta_{kp}\ln(\mathrm{QICAP}) + 2.0\,\beta_{pp}\ln(\mathrm{DP}) + \beta_{pu}\ln(\mathrm{DU}),\\
E_u &= \beta_u + \beta_{lu}\ln(\mathrm{HRS}) + \beta_{ku}\ln(\mathrm{QICAP}) + \beta_{pu}\ln(\mathrm{DP}) + 2.0\,\beta_{uu}\ln(\mathrm{DU}).
\end{aligned}
$$

We can see from the formulas that returns to density and scale at a plant will depend somewhat on the levels of HRS, QICAP, DP and DU during the quarter.

Table 21.1 displays returns to density and scale computed using the sample means for ln(HRS), ln(QICAP), ln(DP) and ln(DU) for the good data in the samples. Increasing returns to density and scale correspond to estimates in Table 21.1 that are greater than one; constant returns correspond to estimates that are equal to one; and, decreasing returns to density and scale correspond to estimates that are less than one. Also shown in Table 21.1 are *t*-values for large sample one-tail tests under the null hypothesis that returns to density and scale are constant. The formula for returns to density is a linear function of the estimated parameters of the translog equations, so its standard deviations can be calculated in the usual way from the asymptotic variance–covariance matrix of the ML estimates.[11] For returns to scale the *t*-values are computed with DP and DU treated as controls. The *t*-values are the differences of the estimates from one divided by their asymptotic standard deviations.

Estimates are shown in Table 21.1 for all of the shape- and process-level production functions that we have fit to samples drawn from MODS. Most of our estimated returns

Table 21.1 Returns to density and scale

Mail shape or process	Returns to density		Returns to scale		Plants ≥ 1		Labor variabilities (%)		
	Estimate	*t*-value	Estimate	*t*-value	Density (%)	Scale (%)	Capital fixed	Fixed input proportions	Expansion path
Shape (first handled pieces, all processes)									
All letters	0.7350	48.114	0.9051	20.537	0.33	15.00	196.35	136.06	103.82
Letters inbound	0.6924	49.096	0.9193	13.901	0.33	14.97	204.63	144.44	105.20
Letters outbound	0.5608	91.469	0.8198	22.743	0.00	5.64	249.72	178.32	143.86
All flats	0.7032	42.561	0.8485	23.892	0.04	1.50	157.86	142.20	147.92
Flats inbound	0.6448	49.349	0.8256	26.406	0.00	0.60	171.78	155.10	139.49
Flats outbound	0.5936	62.087	0.9013	7.261	0.04	13.24	177.58	168.48	172.79
All cancellation	0.3829	93.442	0.9144	6.717	0.00	43.63	566.35	261.15	231.77
All parcels and Priority Mail	0.8128	19.655	0.9307	5.958	0.87	16.85	125.61	123.04	115.80
Letter sorting process (total pieces fed or handled)									
MPBCS total	0.8889	11.747	0.8897	9.588	4.77	10.64	136.10	112.50	104.99
MPBCS inbound	0.8512	18.067	0.8720	10.929	7.50	8.19	132.73	117.49	106.53
MPBCS outbound	0.7744	26.380	0.8908	4.295	9.87	25.43	133.16	129.13	132.07
DBCS total	0.8023	42.418	0.9394	13.797	0.41	14.77	178.84	124.64	233.47
DBCS inbound	0.7648	47.083	0.9238	15.421	1.11	13.86	189.36	130.76	259.87
DBCS outbound	0.6246	78.923	1.0302	−2.736	0.00	62.91	176.37	160.11	193.17
OCR total	0.7285	38.357	0.9429	6.271	2.64	31.22	202.99	137.27	93.89
OCR inbound	0.5677	54.748	0.8432	8.113	0.32	0.59	221.30	176.14	129.43
OCR outbound	0.6508	59.036	0.9361	5.524	2.01	43.38	208.63	153.65	111.11
Manual letters total	0.7377	29.199	0.8024	21.524	0.00	0.00	140.53	135.55	136.34
Manual letters inbound	0.7045	28.175	0.8171	15.052	0.00	0.11	146.59	141.95	141.43
Manual letters outbound	0.6750	29.960	0.8336	12.664	0.04	0.31	164.04	148.14	155.39

Flat sorting process (total pieces fed or handled)									
FSM881 total	0.8741	14.234	0.9292	6.845	4.74	8.80	131.80	114.40	106.40
FSM881 inbound	0.8469	17.228	0.9068	8.432	0.47	3.88	131.69	118.08	112.44
FSM881 outbound	0.8864	13.982	0.9775	1.364	8.98	38.54	120.19	112.81	112.64
FSM1000 total	1.0581	−5.678	0.9296	8.055	68.65	24.74	126.90	94.51	68.18
FSM1000 inbound	0.9553	4.659	0.9263	7.165	37.69	30.01	147.45	104.68	64.74
FSM1000 outbound	0.8354	21.477	0.8863	8.386	5.01	31.53	131.43	119.70	104.37
AFSM100 total	0.8551	23.169	0.9151	14.416	0.29	0.43	124.92	116.95	112.18
AFSM100 inbound	0.8508	20.049	0.9126	11.704	0.81	1.05	127.84	117.53	111.41
AFSM100 outbound	0.6947	40.311	0.9290	3.948	0.14	27.46	126.22	143.95	127.92
Manual flats total	0.9397	6.035	0.8472	15.265	0.00	2.64	109.02	106.42	106.66
Manual flats inbound	0.9225	7.869	0.8500	14.937	7.40	2.44	109.03	108.40	107.71
Manual flats outbound	0.9747	1.401	0.8338	8.848	38.50	0.11	114.44	102.59	104.58
Cancellation process (total pieces handled)									
AFCS total	0.4607	75.787	1.0461	−3.797	0.23	67.07	452.91	217.07	209.49
Non-AFCS	0.2882	50.283	0.5799	12.769	0.00	0.01	440.19	347.04	340.68
Parcels process (total pieces fed or handled)									
SPBS total	0.7875	25.110	0.9695	2.992	0.48	31.84	139.61	126.98	121.28
SPBS other	0.7984	24.730	0.9886	1.020	0.00	42.18	138.57	125.25	119.65
SPBS Priority Mail	0.8162	10.420	1.1882	−4.877	6.76	66.32	122.08	122.52	121.83
Manual parcels	0.5047	19.862	0.4852	17.897	0.06	0.00	204.61	198.13	204.72
Priority Mail	0.8094	8.909	0.8772	4.649	3.57	12.14	120.96	123.54	125.48

to density and scale are less than one at confidence levels that exceed 99 percent. In fact, the majority of the t-values are so high that there is virtually no chance that returns to either density or scale could actually exceed one for any of the shapes and most of the processes at an average USPS mail processing plant.

It is also unlikely that many of the individual plants operate at levels that leave increasing returns. We have calculated returns to density and scale for each plant/quarter in the MODS samples. These have been used to compute the percentages of the plant/quarters for which the returns to density and scale exceed one as weighted averages for the good data using the conditional probabilities. These percentages are displayed for every shape and process in Table 21.1. Only a few exceed 50 percent; most are very small numbers. The USPS processing network as currently configured presents few opportunities to increase productivity by consolidating processing plants to exploit increasing returns to scale.

It is usually the case in Table 21.1 that returns to scale exceed returns to density for a shape or process. This occurs because our econometrics has yielded estimates of E_p and E_u that are mostly positive and statistically significant. The explanation for this result is that ln(DP) and ln(DU) may serve as either output measures, controls or some combination of the two in the production functions. As output measures for a network service activity, we would expect their elasticities to be negative. As controls for an industrial process, the elasticities are more likely to be positive, because ln(DP) and ln(DU) indirectly represent the limitations on continuous operations imposed by the schedule of arrivals and dispatches from the plant. Clearly, our results are best explained by the view that mail processing is primarily an industrial process.

8. LABOR VARIABILITIES

The variabilities that are used to attribute USPS postal costs are the elasiticites of the inputs for various cost categories with respect to single drivers that serve as intermediate output proxies. Mail-processing labor is one of the cost categories and its driver is piece handlings. Therefore, the mail-processing labor variability, ε_l, for a shape is d ln(HRS)/d ln(FHP), and for a process is d ln(HRS)/d ln(TPF). The conventional microeconomic definition for ε_l is that it is the elasticity of the derived demand for labor with respect to output.

The most direct econometric approach to estimating ε_l is to fit the derived demand for labor as a function of piece handlings, relative prices and various controls, and then extract the elasticity from the estimates. This is the route taken by Bozzo (2006), Roberts (2006), Neels (2006) and all others who have submitted econometric estimates of mail-processing labor variabilities derived from the MODS data in postal rate proceedings since 1997.

An alternative approach is to derive the variabilities from econometric estimates of production functions by making one of many possible assumptions about the way that inputs would be adjusted within a processing plant to deal with changes in output. Labor variabilities for three such assumptions are shown in Table 21.1. The assumptions are as follows:

1. *Capital fixed*. Labor variability is derived under the assumption that HRS is variable but QICAP is fixed. This assumption best describes the operation of USPS mail-processing plants in the very short run: $\varepsilon_l = 1/E_l$.

2. *Fixed input proportions*. Labor variability is derived under the assumption that HRS/QICAP is fixed; that is, HRS and QICAP respond proportionately to changes in FHP or TPF. The variabilities of labor and capital are the same: $\varepsilon_l = \varepsilon_k$. This assumption corresponds to the way that USPS and the PRC use the labor variabilities in postal cost accounting. ε_l is regarded as a constant and is applied to both labor costs and the associated mail-processing equipment cost categories. The practice is called 'piggy backing'. The assumption can be defended as an approximation to USPS operating plans in the intermediate run. USPS does not rethink the way that it equips and staffs its processing plants every time postal volumes change. Instead, USPS rescales the existing levels of labor and equipment to handle such changes without altering the input proportions. Reevaluations of the proportions are likely to occur much less frequently and to be initiated by changes in equipment technology and relative factor prices rather than volumes: $\varepsilon_l = 1/(E_l + E_k)$.
3. *Expansion path*. Labor variability is derived under the assumption that HRS and QICAP are maintained in an economically efficient combination at the relative factor prices implied by the first-order conditions for minimizing cost under the production function. This assumption corresponds to a long-run equilibrium in which both HRS and QICAP are variable and USPS optimally adjusts the input levels at its plants. Under this assumption, HRS and QICAP move in tandem along an expansion path that is determined by a fixed ratio of factor prices. The variability of labor, ε_l, the variability of capital, ε_k, and the elasticity of a Lagrange multiplier, $\varepsilon_\lambda = d \ln(\lambda)/d \ln(\text{FHP})$, are calculated by solving a system of three simultaneous linear equations:

$$(2.0\beta_{ll} - E_l)\varepsilon_l + \beta_{lk}\varepsilon_k + E_l\varepsilon_\lambda = 0,$$
$$\beta_{lk}\varepsilon_l + (2.0\beta_{kk} - E_k)\varepsilon_k + E_k\varepsilon_\lambda = 0,$$
$$E_l\varepsilon_l + E_k\varepsilon_k = 1.$$ [12]

All but four of the mail-processing labor variabilities in Table 21.1 exceed 100 percent, as would be expected for production functions exhibiting decreasing returns to density. There is also a general tendency for the variabilities to progress from highest to lowest in order from short to long run. However, the long-run expansion path variabilities still remain mostly above 100 percent.

9. THE FITS OF THE SWITCHING REGRESSIONS

There are large differences in labor and capital productivity among the mail-processing plants covered by MODS. These differences occur with every shape of mail and for every process. At first glance, they seem hard to explain. The plants all use the same types of processing machines and other equipment. They are organized and staffed following the same USPS guidelines. And, they are all managed according to the same practices and standards. Our ML estimates of the translog production functions largely solve this puzzle. Selections of shape-level and process-level ML estimation results for the stochastic switching regressions are provided in Tables 21.2 and 21.3.

Table 21.2 Selected shape level maximum likelihood regressions

Equation	All letters		All flats		All cancellation		All parcels and Priority Mail	
Dependent variable	ln(FHP) total sorting		ln(FHP) total sorting		ln(TPH)		ln(FHP) total sorting	
Capital index	QIAHE		QIMHE		QIAFCS + QICANC		QIMHE	
Goodness of fit	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n
Adjusted R-square	0.9789	0.8282	0.9747	0.9282	0.9553	0.8331	0.9382	0.7833
Standard deviation of error	0.1241	0.6049	0.1363	0.3249	0.1736	0.7546	0.2757	0.8180
Number of observations	8435	136	6686	1887	7994	325	7013	1440
	Percent	t-value	Percent	t-value	Percent	t-value	Percent	t-value
Bad data	1.59	9.548	22.01	20.854	3.90	12.285	17.03	22.718
	Good equation		Good equation		Good equation		Good equation	
Control variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
Intercept	11.9023	24.279	11.9291	13.631	2.6502	3.208	0.3044	0.197
Quarter 1	0.0364	9.236	0.1266	21.275	0.0232	4.068	0.0237	2.111
Quarter 2	0.0484	12.047	0.0077	1.325	0.1932	33.079	0.0015	0.134
Quarter 3	0.0352	9.061	0.0927	15.933	0.0658	11.519	0.0322	2.948
Annual trend	0.0168	17.109	0.0220	11.907	−0.0294	−19.292	0.0260	10.324
CURB share	0.0752	2.944	0.1867	4.731	−0.1021	−2.552	0.4421	5.704
NDCBU share	0.6542	17.443	0.4963	9.366	0.4186	7.295	0.1058	1.008
CENT share	−0.3432	−14.929	0.0742	1.744	0.2231	5.535	−0.3847	−4.923
RB share	−0.0040	−0.260	−0.0798	−3.633	−0.1023	−4.516	−0.2238	−5.356
HCT share	0.2146	2.885	0.1403	1.360	0.2616	2.574	1.1493	6.396
POBOX share	−0.3615	−11.358	−0.3621	−7.451	−1.1180	−22.465	−0.1377	−1.632
1st Presort share	0.0225	0.909	−0.4121	−2.421	−0.3060	−14.483		
Std CR share	0.1630	3.235	0.0115	0.274	−0.2228	−4.444		
Std Presort share	0.1342	5.290	0.0402	0.784	−0.1423	−5.808	−0.3036	−4.276
Periodicals share			0.0806	1.107			−0.1471	−1.875
Parcel Post share							−0.3041	−3.025

BPM share							−0.3388	−3.726
Media Mail share							0.4828	3.786
ln (origin/destination)	0.0632	19.351	0.0521	13.603	0.1582	23.425	0.0642	8.155
No parcels	0.0164	1.507	0.0201	0.900	−0.0334	−2.952		
No Priority Mail					0.0234	4.316	−0.1108	−8.206
No cancellation	0.0737	9.150	−0.0040	−0.288			0.1912	7.413
No MPBCS	−0.0097	−2.313			0.0190	3.316		
No DBCS	−0.5929	−22.193						
No OCR	−0.0839	−12.139			−0.0606	−6.252		
No AFCS	−0.0395	−5.839			−0.1377	−11.445		
No FSM881			0.0893	13.744				
No FSM1000			0.0548	9.592				
No AFSM100			−0.1554	−22.636				
No SPBS Other							−0.0586	−4.509
No SPBS Priority							−0.0448	−3.822
Auto labor share	0.4738	12.606	1.1305	16.951	0.5768	8.843		
Mechanized labor share							−1.6617	−9.741
Manual labor share								
Other labor share	0.5404	21.429	0.9862	25.492	0.0445	1.112	−0.0570	−0.854
Elasticity at mean	Estimate	*t*-value	Estimate	*t*-value	Estimate	*t*-value	Estimate	*t*-value
Labor hours	0.5093	93.398	0.6335	86.894	0.1766	38.936	0.7961	96.575
Capital index	0.2257	36.766	0.0698	16.444	0.2063	38.681	0.0166	2.213
Delivery points	0.1237	18.871	0.1106	11.776	0.4649	50.450	0.0772	5.101
Delivery units	0.0643	13.537	0.0605	8.234	0.1163	16.358	0.0495	3.602

Table 21.3 Selected process level maximum likelihood regressions

Equation	DBCS total		Manual letters total		AFSM100 total		SPBS Other	
Dependent variable	ln(TPF)		ln(TPH)		ln(TPF)		ln(TPF)	
Capital index	QIDBCS		QIPSE		QIAFSM		QIMHE residual	
Goodness of fit	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n	Good eq'n	Bad eq'n
Adjusted R-square	0.9790	0.8276	0.9299	0.8235	0.9802	0.9741	0.9422	0.8019
Standard deviation of error	0.1306	0.5805	0.2398	0.4390	0.0906	0.1102	0.2457	0.8568
Number of observations	8215	305	7353	1134	2776	395	5345	209
	Percent	t-value	Percent	t-value	Percent	t-value	Percent	t-value
Bad data	3.58	10.618	5.82	10.007	12.46	7.763	3.76	9.150
Control variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
Intercept	2.7142	4.870	8.7399	7.667	3.0581	3.508	2.0618	1.196
Quarter 1	0.0291	6.756	0.0511	6.171	0.0667	11.507	0.0634	6.136
Quarter 2	0.0576	13.180	0.0213	2.521	0.0077	1.367	0.0164	1.608
Quarter 3	0.0267	6.242	0.0480	5.921	0.0408	7.235	0.0447	4.482
Annual trend	0.0300	28.790	−0.0206	−8.902	0.0344	11.813	0.0022	0.884
CURB share	0.2117	6.969	0.0995	1.793	0.0537	1.453	−0.0335	−0.456
NDCBU share	0.7591	18.393	0.9695	13.647	0.1921	4.228	0.0452	0.494
CENT share	0.1352	5.110	−0.3596	−6.572	0.0995	2.842	−0.4440	−6.142
RB share	0.0688	4.196	−0.1919	−6.353	−0.0312	−1.548	−0.2489	−6.511
HCT share	−0.3942	−5.415	−0.1941	−1.377	0.0955	0.786	−0.7973	−4.281
POBOX share	−0.2627	−7.441	−0.3929	−5.853	−0.0697	−1.196	0.2407	2.623
1st Presort share	0.1799	6.820	0.2526	4.901	−0.2474	−1.516		
Std CR share	0.0838	1.448	0.1436	1.273	−0.1471	−3.236		
Std Presort share	0.1955	6.997	0.2774	4.693	−0.0225	−0.396	−0.4769	−7.152
Periodicals share					−0.1401	−1.698	−0.5710	−7.649
Parcel Post share							−0.3242	−3.415
BPM share							−0.4608	−4.828
Media Mail share							−0.0527	−0.382

ln (origin/ destination)	0.0490	13.178	−0.0026	−0.340	0.0191	4.279	0.0832	11.487
No parcels	0.0166	1.807	−0.0825	−2.551	−0.0232	−1.742		
No Priority Mail	−0.0162	−3.985			−0.0159	−2.581	0.1071	9.204
No cancellation	0.0098	1.059	0.0325	1.820	0.0132	0.774	0.1420	4.598
No MPBCS	0.0179	3.958	−0.0424	−4.835				
No DBCS			−0.4981	−11.278				
No OCR	−0.0835	−10.686	−0.0645	−4.514				
No AFCS	−0.0923	−12.533	−0.1015	−6.946				
No FSM881					0.0205	2.428		
No FSM1000					−0.0186	−3.231		
No AFSM100								
No SPBS Other								
No SPBS Priority							0.0805	8.771
Auto labor share	−0.6830	−15.335			0.0616	0.903		
Mechanized labor share							−0.9772	−6.372
Manual labor share			0.7966	15.661				
Other labor share	−0.1292	−4.741	0.3897	7.028	0.4499	9.987	0.2836	4.133
Elasticity at mean	Estimate	*t*-value	Estimate	*t*-value	Estimate	*t*-value	Estimate	*t*-value
Labor hours	0.5592	121.408	0.7116	81.054	0.8005	123.490	0.7216	88.336
Capital index	0.2432	52.354	0.0262	3.745	0.0546	10.098	0.0767	14.391
Delivery points	0.0958	15.678	0.0287	2.166	0.0749	8.652	0.2142	15.150
Delivery units	0.0502	10.066	0.0519	5.333	−0.0094	−1.227	−0.0218	−1.601

One of the explanations to be found in the tables is that there can be a lot of bad data in the MODS samples. The percentage of bad data found by the ML estimator ranges from about 1.5 percent (letters) to over 20 percent (flats). The high t-values for these percentages mean that there is little chance that the MODS samples could actually be clean. Much of the bad data consists of observations of extreme values for FHP (or TPF), HRS or QICAP. Such observations are assigned high conditional probabilities of being bad by the ML estimator. Observations with high conditional probabilities also appear mostly in runs and clusters associated with single plants.

Our translog model fits the MODS samples quite well despite the bad data. The standard deviation of the errors and adjusted R-squares for the good and bad equations have been corrected for degrees of freedom as is customary for an OLS regression. The R-squares for the good equations are mostly in the neighborhood of 0.89 to 0.98. The R-squares for the bad equations are somewhat lower, but still high enough to show in every case that the bad data is explained well by a distinct linear regime. The standard errors of the good equations are also all less than the standard errors for the bad equations. This accords well with the kinds of dirty samples analyzed in Pearsall (2007). Notice, however, that the standard errors for the good and bad equations for the AFSM 100s in Table 21.3 are almost identical. This has happened because many of the observations given high conditional probabilities of being bad correspond to periods when the new AFSM 100s were first installed in the plants. The bad equation is actually a fit of the production function for a startup regime.

Coefficient estimates and t-values are shown for the good equations in Tables 21.2 and 21.3. Many of the t-values in every equation are far above (absolutely) the critical values for two-tailed tests at the 95 and 99 percent confidence levels. The reason for this high precision in the good data estimates is that the MODS samples are so large that the ML estimator can accurately separate the effects of the good and bad data. The overall conclusions to be drawn from the estimates are, first, that the variables and controls included in the production functions as explanatory variables are statistically effective, and, second, that a two-regime stochastic switching regression is a good model of the flawed data generation process that produced the dirty MODS samples.

10. OTHER DETERMINANTS OF PRODUCTIVITY

Returns to density and scale are only one of many statistically significant determinants of productivity. There are high t-values for one or more of the variables in every group of controls in almost every equation. This is somewhat surprising, since many of the controls are no better than indirect proxies for plant characteristics that USPS will not disclose. This means that there are often multiple explanations for the signs and magnitudes of the coefficient estimates. Nevertheless, the estimates exhibit patterns that support several hypotheses regarding the determinants of mail-processing productivity besides returns to density and scale.

There is a seasonal pattern to productivity. USPS processing plants are less productive during postal quarter 4 (July to September) than at other times of the year. Most regular employees take their annual vacations during this quarter. The drop in productivity could also be explained by seasonal variations in weight and other hedonic properties of the average piece.

Productivity for most shapes and processes exhibits a highly significant autonomous trend. For automated and mechanized processes the trend is usually positive; for manual processes it is negative. These patterns could be the result of frequent minor upgrades to processing equipment and a deterioration of labor skills as manual sorting is slowly phased out. The increasing use of automated processes could also be leaving more difficult residues of pieces to be processed by hand.

The delivery point shares are proportions of DP, a variable that indirectly controls for arrival and dispatch schedule effects on productivity. The coefficient estimates for the delivery point shares follow a distinct pattern confirming that schedule effects are among the most important determinants of productivity. City central deliveries (CENT) and P.O. boxes (POBOX) are made early in the day, while other deliveries are often made later. Deliveries to rural boxes (RB) take longer than city and suburban deliveries, so the mail must be available to rural carriers at delivery units early in the day. A plant with high percentages of CENT, RB and POBOX delivery points in its service territory is likely to have a more difficult schedule for dispatching processed incoming mail to its delivery units. This explains the evident tendency for the coefficients of these shares to have a negative sign. On the other hand, mail delivered to central box units (NDCBU) is usually delivery point sorted only to the box unit. Such mail requires fewer handlings and appears to be somewhat easier to process than other mail. The estimated coefficients for the NDCBU share of delivery points are mostly positive numbers.

For the shape-level production functions in Table 21.2 we would expect presorted mail to increase productivity because it requires fewer piece handlings to process. For the process-level functions in Table 21.3 we would also expect presorted mail to increase productivity, but for a different reason. Presorted mail conforms to a higher standard of address hygiene than other mail. This should make this mail easier to process using automated equipment. These effects seem to be present in the coefficients for letters, flats, cancellations and the processes for letters and flats, except for the AFSM 100. For parcels, the coefficients of the subclass and worksharing shares probably are a reflection of the effect on processing of differences in the average sizes and weights of parcels in the categories. First-class parcels, priority parcels and media mail all tend to be smaller and lighter than other kinds of parcels.

Outgoing (originating) mail usually requires less processing than incoming (destinating) mail because it is sorted to fewer divisions. The coefficients of ln(origin/destination) in Table 21.2 are all positive and highly significant. However, this cannot be the only effect represented by the variable 'ln(origin/destination)'. The coefficient is also positive and significant in most of the shape-level equations that distinguish between incoming and outgoing pieces, and in most of the non-manual process-level equations with ln(TPF) as the output variable, including those shown in Table 21.3. The positive coefficients in these other equations show that a higher proportion of outgoing (originating) mail for a shape makes all of the automated and mechanized processes for the shape more productive. It is not really clear why this occurs. It may be that large mailers, whose mail is generally easier to process, prefer to submit their mailings at the more productive plants, or that USPS caters to large mailers by making the plants they use more productive.

The shape dummy variables for parcels, priority and cancellations were included to indirectly identify the type of plant. Significant coefficients with both positive and negative values can be found frequently among our estimates of these coefficients, so the type of plant is an important determinant of productivity.

The process dummy variables in each of the production functions identify the equipment present in the plant that pertains to the associated mail shape. Most of the estimated coefficients for these dummy variables turn out to be statistically significant at very high levels, including all of those shown in Tables 21.2 and 21.3. The presence or absence of complementary shape-specific equipment appears to be the single most important determinant of productivity at USPS processing plants.

The coefficients of the process dummies often display a distinctive pattern. The largest negative coefficient is frequently associated with the most modern automated type of equipment. The coefficients then grow smaller absolutely as they progress through the equipment types, and sometimes turn positive as the oldest and least automated types are reached. We can observe this pattern in many of the fits shown in the tables. For example, the coefficient for the variable 'No DBCS' is negative and is the largest absolutely in the equations for 'All Letters' and 'Manual Letters'. Next in order are the negative coefficients for 'No OCR' and 'No AFCS'. The coefficient for 'No MPBS', the oldest of the letter-processing machines, is still negative but is smaller than the others.

The labor skill shares are the least satisfactory of the controls included in the equations. They are included because the MODS data offer no other choice for representing the skill levels of a plant's labor force. The shares actually represent work assignments that may or may not be based on skills. The coefficient estimates are usually significant, but difficult to interpret as evidence of the effects of specific labor skills on productivity.

The elasticities of FHP (or TPF) with respect to HRS, QICAP, DP and DU have been computed from the formulas given earlier and are displayed at the bottom of Table 21.2 and Table 21.3 for the selected production functions.[13] All of the estimated elasticities for labor, E_l, are greater than zero, but less than one as would be expected for an input subject to diminishing returns. The elasticities for capital, E_k, are also in the zero-to-one range with a few process-level exceptions, of which only one (AFSM 100 Outbound) is statistically significant. The capital elasticities are typically much smaller than the labor elasticities, especially for the manual processes. All cancellations and cancellation using the AFCSs are an important exception in which the labor and capital elasticities are almost equal. As we have already noted, the estimates of E_p and E_u in Tables 21.2 and 21.3 are mostly positive and statistically significant.

11. CONCLUSION

Our stochastic switching models' ML estimates supply a statistically robust answer to the question in our title. USPS mail processing is commonly conducted in its plants at volume levels that are sufficiently high to encounter decreasing returns to density and scale. We have found generally decreasing returns for plant-level aggregates of pieces handled by shape, for pieces fed in single processes, and for most inbound and outbound sub-streams of the mail.

The finding that the average plant operates in the region of decreasing returns to density translates under several assumptions into average variabilities for labor in mail processing that exceed 100 percent. Therefore, the use of variabilities that are less than 100 percent in USPS cost accounting will lead to under-estimates of the marginal costs and attributable costs of processing the mail.

The finding that individual processing plants mostly operate in the region of decreasing returns to scale means that simply consolidating plants is not likely to be an effective strategy for restructuring the USPS network with the object of increasing aggregate productivity. Most plant consolidations will actually decrease the volume that can be processed by the same equipment and labor force in the consolidated plants.

Finally, there are strong indications in our estimates that factors other than scale are chiefly responsible for the large observable differences in average productivities among USPS plants. Perhaps the most interesting of these indications is the indirect evidence we have found that productivity is affected by a plant's ability to schedule arrivals and dispatches to smooth mail flows through its processing operations.

NOTES

* The views expressed in this chapter are those of the authors and do not necessarily represent the opinions of the US Postal Regulatory Commission.

1. For a more detailed description of USPS mail processing, see Bozzo (2006, pp. 11–33).
2. In the cost studies increasing returns are evidenced by decreasing marginal costs and vice versa. Merewitz (1971), Gupta (1982), Strack (1986), Moriarty et al. (2006), Cohen and Chu (1997), Roberts (2006), and Neels (2006), have all found decreasing returns at some level of mail processing and/or overall postal operations. On the other hand, studies by Panzar (1984), Kleindorfer (1987), Norsworthy et al. (1991), Bradley and Colvin (1999), Bozzo (2006), Wells (1987), and Rogerson and Takis (1993) have found increasing returns.
3. See Bozzo (2006). The samples were constructed by combining three separate worksheets found in Library References sponsored by Bozzo for the R2006-1 omnibus rate proceeding. The Library References accompanied his direct testimony on behalf of USPS and his responses to interrogatories and requests for data from United Parcel Service and the PRC's Office of the Consumer Advocate.
4. Bozzo (2006) and Roberts (2006) fit derived demand functions for labor to the MODS data under the assumption that capital is fixed. This is clearly not a desirable underlying assumption for a study of returns to density and scale. Bozzo and Roberts' models also explicitly rely on an assumption that USPS mail processing is both economically and technologically efficient. Fitting production functions, as we have done, enlists only the assumption that processing is technologically efficient.
5. There is an element of moral hazard in the MODS self-collection system that may cause malfunctions. The MODS reports are used by USPS's higher management partly to assess the performance of the plants.
6. See, especially, Neels (2006) on the specific defects in the screens proposed by Bozzo (2006).
7. The ML estimator is described in detail in a manuscript that is available from one of the authors. See Pearsall (2007). ML estimation of stochastic switching models was first proposed by Quandt (1972) and extended by Goldfeld and Quandt (1973), Maddala and Nelson (1975), Quandt and Ramsey (1978), Hartley (1978), and Hamilton (1989). Exogenous switching regression models have been widely applied and have been particularly useful in econometric studies of markets in disequilibrium. ML estimation is the method of choice for fitting stochastic switching models with exogenous fixed-probability independent switching. It is the method described in Judge et al. (1985) and implemented in statistical software such as LIMDEP and SAS.
8. Several of these worksheets are available at www.prc.gov.
9. Our practice was to choose the capital equipment index that most closely matched the shape or process for each equation. Manual processes were fit using the capital index for postal support equipment (QIPSE). The capital index for SPBS was calculated as a residual by subtracting all other mechanized process indices from the index for total mechanized processing equipment (QIMPE). The capital service indices for the incoming and outgoing equations were pre-multiplied by the proportions of incoming and outgoing work hours.
10. The derivation of the formula for returns to scale is as follows:

    $t^k y = f(t^k P, t^k U, tL, tK)$ the production function with y = FHP, TPF (or TPH), P = DP, U = DU, L = HRS and K = QICAP; k is the measure of returns to scale and t is a scale factor such that $t=1$ gives an input/output combination on the production function.

    $kt^{k-1}y = kt^{k-1}f_p P + kt^{k-1}f_u U + f_l L + f_k K$ differentiate through with respect to t. f_p, f_u, f_l and f_k are the partials of f with respect to P, U, L and K.

$ky = kf_pP + kf_uU + f_lL + f_kK$ the equation must hold at $t = 1$.
$k = (f_lL + f_kK)/(y - f_pP - f_uU)$ solve for k.
$k = (f_lL/y + f_kK/y)/(1 - f_pP/y - f_uU/y)$ divide the numerator and denominator by y.
$k = (E_l + E_k)/(1 - E_p - E_u)$ substitute the elasticities in the formula for k.

A measure of returns to scale that treats DP and DU as controls is $E_l + E_k + E_p + E_u$. Statistical tests under the null hypothesis (H_0) that returns to scale equal one yield identical results for this measure and for the measure with DP and DU treated as outputs. However, $E_l + E_k + E_p + E_u$ is linear in the parameter estimates of the translog production function, so its standard deviation can be computed readily from the variance–covariance matrix of the estimates. The t-values for returns to scale shown in Table 21.1 are computed for $H_0: E_l + E_k + E_p + E_u = 1$.

11. The calculation of the asymptotic variance–covariance matrix follows Maddala (1977, pp. 176–81) and Judge et al. (1985, pp. 177–80). Standard deviations are calculated for linear combinations of the model parameters. A complete description is in Pearsall (2007).
12. These equations are derived from the first-order conditions for a cost minimization with the translog production function as a constraint and the price ratio P_k/P_l held constant. The translog production function can be reduced to the terms involving ln(HRS) and ln(QICAP) for this derivation:

$$\begin{aligned}\ln(\text{FHP}) = {} & [\beta_1 + \beta_{lp}\ln(\text{DP}) + \beta_{lu}\ln(\text{DU})]\ln(\text{HRS}) \\ & + [\beta_k + \beta_{kp}\ln(\text{DP}) + \beta_{ku}\ln(\text{DU})]\ln(\text{QICAP}) \\ & + \beta_{ll}\ln(\text{HRS})^2 + \beta_{lk}\ln(\text{HRS})\ln(\text{QICAP}) + \beta_{kk}\ln(\text{QICAP})^2 + \text{other terms.}\end{aligned}$$

Using the elasticities for HRS and QICAP, the first-order conditions are as follows:

$$\text{HRS} + \lambda E_l = 0,\ (P_k/P_l)\text{QICAP} + \lambda E_k = 0, \text{ and the translog production function.}$$

With ln(HRS), ln(QICAP), λ and ln(FHP) as variables, we take total differentials of the first-order conditions with P_k/P_l held fixed:

$$\begin{aligned}&\text{HRS d}\ln(\text{HRS}) + E_l\text{d}\lambda + \lambda[2.0\beta_{ll}\text{d}\ln(\text{HRS}) + \beta_{lk}\text{d}\ln(\text{QICAP})] = 0\\ &(P_k/P_l)\text{QICAP}\ln(\text{QICAP}) + E_k\text{d}\lambda + \lambda[\beta_{lk}\text{d}\ln(\text{HRS}) + 2.0\beta_{kk}\text{d}\ln(\text{QICAP})] = 0\\ &E_l\text{d}\ln(\text{HRS}) + E_k\text{d}\ln(\text{QICAP}) = \text{d}\ln(\text{FHP}).\end{aligned}$$

Next, we substitute for HRS and (P_k/P_l)QICAP from the first-order conditions and divide the first two differentials by λ. After rearranging terms, the differentials become:

$$\begin{aligned}&(2.0\beta_{ll} - E_l)\text{d}\ln(\text{HRS}) + \beta_{lk}\text{d}\ln(\text{QICAP}) + E_l(\text{d}\lambda/\lambda) = 0,\\ &\beta_{lk}\text{d}\ln(\text{HRS}) + (2.0\beta_{kk} - E_k)\text{d}\ln(\text{QICAP}) + E_k(\text{d}\lambda/\lambda) = 0,\\ &E_l\text{d}\ln(\text{HRS}) + E_k\text{d}\ln(\text{QICAP}) = \text{d}\ln(\text{FHP}).\end{aligned}$$

We divide all three differentials through by d ln (FHP). The equations in the text result from the substitutions ε_l = d ln(HRS)/d ln(FHP), ε_k = d ln(QICAP)/d ln(FHP), and ε_λ = d ln(λ)/*d* ln(FHP) = (dλ/λ)/d ln(FHP).
13. The 14 'beta' coefficients of the translog segments of the production functions do not have individual economic meanings. However, these coefficients collectively determine the elasticities. The calculations used the sample means of ln(HRS), ln(QICAP), ln(DP) and ln(DU) for the good data. Standard deviations were derived from the asymptotic variance–covariance matrices for the ML estimator and used to compute the t-values for the elasticities.

REFERENCES

Bozzo, A.T. (2006), Direct Testimony (USPS-T-12), sponsored Library References and Responses to Interrogatories, PRC Docket No. R2006-1.

Bradley, M.D. and J. Colvin (1999), 'Productivity and technical change in a public service enterprise,' in M.A. Crew and P.R. Kleindorfer (eds), *Emerging Competition in Postal and Delivery Services*, Boston, MA: Kluwer Academic Publishers, pp. 75–88.

Cohen, R.H. and E. Chu (1997), 'A measure of scale economies in postal services,' in M.A. Crew and P.R. Kleindorfer (eds), *Managing Change in the Postal and Delivery Industries*, Boston, MA: Kluwer Academic Publishers, pp. 115–32.

Goldfeld, S.M. and R.E. Quandt (1973), 'A Markov model for switching regressions', *Journal of Econometrics*, **1**, 3–16.
Gupta, S.N. (1982), 'Production and cost functions of the United States Postal Service', Ph.D. dissertation, The George Washington University.
Hamilton, J.D. (1989), 'A new approach to the economic analysis of non-stationary time series and the business cycle', *Econometrica*, **57**, 357–84.
Hartley, M. (1978), 'Comment' (on Quandt and Ramsey), *Journal of the American Statistical Association*, **73**, 738–41.
Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T.-C. Lee (1985), *The Theory and Practice of Econometrics*, New York: John Wiley and Sons, pp. 177–80, 806.
Kleindorfer, P.R. (1987), Direct Testimony (USPS-T-4), PRC Docket No. R87-1.
Maddala, G.S. (1977), *Econometrics*, New York: McGraw-Hill, pp. 176–81.
Maddala, G.S. and F.D. Nelson (1975), 'Switching regression models with exogenous and endogenous switching', *American Statistical Association Proceedings of Business and Economics Statistics Section*, pp. 423–6.
Merewitz, L. (1971), 'Costs and returns to scale in U.S. post offices,' *Journal of the American Statistical Association*, **66**, 504–9.
Moriarty, R., S. Yorke, G. Harman, J. Cubbin, M. Meschi and P. Smith (2006), 'Economics analysis of the efficiency of Royal Mail units and the implications for regulatory policy', in M.A. Crew and P.R. Kleindorfer (eds), *Liberalization of the Postal and Delivery Sector*, Cheltenham, UK and Northampton, MA, USA, Edward Elgar, pp. 165–82.
Neels, K. (2006), Direct Testimony (UPS-T-2), Docket No. R2006-1.
Norsworthy, J.R., S.-L. Jang and W.-M. Shi (1991), 'Productivity and cost measurement for the United States Postal Service: variations among regions', in M.A. Crew and P.R. Kleindorfer (eds), *Competition and Innovation in Postal Services*, Boston, MA: Kluwer Academic, pp. 141–72.
Panzar, J.C. (1984), Rebuttal Testimony on Behalf of ANPA, PRC Docket No. R84-1.
Pearsall, E.S. (2007), 'A dirty data estimator', manuscript available at www.prc.com.
Quandt, R.E. (1972), 'A new approach to estimating switching regressions', *Journal of the American Statistical Association*, **67**, 306–10.
Quandt, R.E. and J.M. Ramsey (1978), 'Estimating mixtures of normal distributions and switching regressions', *Journal of the American Statistical Association*, **73**, 730–8.
Roberts, M. (2006), Direct Testimony (OCA-T-1), Docket No. R2006-1.
Rogerson, C. and W. Takis (1993), 'Economies of scale and scope and competition in postal services', in M.A. Crew and P.R. Kleindorfer (eds), *Regulation and the Nature of Postal and Delivery Services*, Boston, MA: Kluwer Academic Publishers, pp. 109–27.
Strack, W.D. (1986), 'Returns to scale and economies of scope in large mail processing facilities,' Rate Research Division, Office of Rates, USPS.
US Postal Rate Commission (1998, 2001, 2002 and 2006), Postal Rate and Fee Changes, Opinion and Recommended Decision, Cost Attribution Section and Appendices F and I, Docket Nos. R97-1, R2000-1, R2001-1 and R2005-1.
US Postal Rate Commission (2006), Evolutionary Network Development Service Changes Advisory Opinion Concerning a Proposed Change in the Nature of Postal Services, Docket No. N2006-1
US Postal Regulatory Commission (2007), Postal Rate and Fee Changes, Opinion and Recommended Decision, Mail Processing Variability Section and Appendix J, Docket No. R2006-1.
Wells, F.J. (1987), Direct Testimony (USPS-T-5), PRC Docket No. R87-1.
Wilkes, S.S. (1962), *Mathematical Statistics*, New York: John Wiley and Sons, pp. 344–65.