USPS-LR-J-56

Programs and Electronic Input Data for Mail Processing Volume Variability Analysis (USPS-T-14)

Category Two Library Reference
Witness Bozzo

USPS-LR-J-56

Programs and Electronic Input Data for Mail Processing Volume Variability Analysis

(USPS-T-14)

Table of Contents

Introduction

This library reference contains background material for the econometric analysis reported in the testimony of witness Bozzo (USPS-T-14). It fulfills the same purpose in this proceeding as did USPS-LR-I-107 in the last case. Three main sections follow this introduction. Section I contains descriptions of the programs used to estimate the recommended volume-variability factors, as well as the various alternative estimators and specification tests described in the testimony. Section II contains descriptions of the computer programs and processing procedures used to assemble the data set used in the estimation procedures. Finally, Section III contains a description of the methods used to develop MODS productivity data for use by witnesses Miller (USPS-T-22 and USPS–T–24). The CD-ROM that accompanies this library reference contains electronic versions of all computer programs used to produce results found in USPS-T-14, the full econometric outputs derived from them, and the input data used.

I. Programs Used to Estimate Mail Processing Volume-Variability Factors and Related Statistics

All econometric results presented in USPS-T-14 were produced with TSP Version 4.4 econometric software[1], running on a 1000-MHz AMD Athlon-based PC with 256 Mbytes of RAM and the Windows 2000 operating system.[2] This section describes each program—its purpose, the main data inputs it uses, the main outputs it produces, and its logical flow. An example program—the main automated/mechanized letter and flat operations program—and its output listing is provided in the Appendix that follows section IV; all programs, listing files, and data are provided on the accompanying CD-ROM. All TSP programs have the file extension ".tsp"; the corresponding output listings have the same file names with the file extension ".out".[3] In the descriptions and definitions that follow, <op> is a placeholder denoting a MODS operations group, and <ldc> is a placeholder denoting the NWHRS Labor Distribution Code (LDC) corresponding to a MODS operations group.

A. Automated/Mechanized Letter and Flat Distribution Operations Program

The program "varmp-tpf-by2000.tsp" produces the recommended volume-variability factors for the BCS (BCS/ and BCS/DBCS), OCR, FSM (FSM/ and

---

[1] See B.H. Hall and C. Cummins, *Time Series Processor Version 4.4 Reference Manual*, TSP International March 1998; and B.H. Hall and C. Cummins, *Time Series Processor Version 4.4 User's Guide Including an Introductory Guide*, TSP International August 1997.

[2] I also replicated the main results of the TSP programs using SAS release 8.00 statistical software, running on a 733-MHz Intel Pentium III-based PC with 256 Mbytes of RAM and the Windows 2000 operating system.

FSM/1000), LSM, and SPBS cost pools, together with associated statistics. This subsection describes the program, walking the reader through each section of the program in order and discussing the major variables. The subsequent programs, which are based on varmp-tpf-by2000.tsp, and have essentially identical variable names and logic flows, will only be described additionally to the extent that they differ from this one.

1. Preliminaries, data extraction, and global data transformations

The initial section of the program sets parameter values and creates lists of variable names to facilitate the reading of the raw data file. The data, which are stored in the Microsoft Excel version 4 workbook file "reg9300.xls", are then read. The following variables are read:

Variable Name	Variable Description
• IDNUM	Unique identification number for each mail processing facility
• QTR	String combining Postal quarter (first digit) and Postal fiscal year (second and third digits)
• TPH<op>	Total pieces handled in operation group <op> for the current Postal quarter
• TPF<op>	Total pieces fed in operation group <op> for the current Postal quarter
• HRS<op>	Total clerk/mailhandler labor hours recorded in operation group <op> for the current Postal quarter
• CURB	Number of city curb delivery points in facility service territory
• NDCBU	Number of city NDBCU delivery points in facility service territory
• CENT	Number of city central delivery points in facility service territory
• OTHER	Number of other city delivery points in facility service territory

[3] All SAS programs have the file extension ".sas"; the corresponding output listings have the same file names with the file extension ".lst"

Variable Name	Variable Description
• RB	Number of rural delivery points in facility service territory
• HCT	Number of highway contract route delivery points in facility service territory
• POBOX	Number of possible P.O. boxes in facility service territory
• LGPO	Number of large post offices in facility service territory
• SMPO	Number of small post offices in facility service territory
• SB	Number of stations and branches in facility service territory
• HOURS<ldc>	Regular plus overtime workhours in LDC <ldc> for the current Postal quarter
• DOLLARS<ldc>	Regular plus overtime compensation in LDC <ldc> for the current Postal quarter
• QIAHE	Quantity index for automated letter sorting equipment (BCS and OCR) capital
• QIMHE	Quantity index for mechanized mail handling equipment capital
• QIPSE	Quantity index for postal support equipment capital
• QIBLD	Quantity index for total buildings capital
• QIPDBLD	Quantity index for mail processing and distribution buildings capital
• QICAP	Quantity index for total facility capital
• SQFT	Facility square footage
• N5DZIP	Number of 5-digit ZIP Codes in facility service territory

Several data transformations are then performed to produce variables needed for the econometric analysis, including total possible deliveries (the sum of the delivery point variables listed above), the time trend, and quarterly dummy variables. The fourth Postal quarter (PQ4) contains four accounting periods rather than three, so the PQ4 observations of piece handlings, hours, and capital input are reweighted to make them comparable with the corresponding values in other quarters. The LDC wages and the

manual ratios are calculated and mapped to MODS operation group, and operation-specific productivity filter bounds are set.

2. Sample Selection

The remainder of the program consists of a logical "loop". Statements that lie between the "dot..." statement that opens this loop and the "enddot" statement that closes it are repeated once for each of the MODS operation groups listed in the "dot..." statement.[4] The next section of the program (the first inside the loop) selects the regression samples. After identifying "all usable" observations (i.e., those with nonmissing positive values for each of the analysis variables, beginning from the first quarter of PFY 1995), filters are applied to eliminate cases with:

- fewer than 40 work hours per quarter recorded in the operation,
- implausibly low productivity (TPF/hr) in the operation, or
- implausibly high productivity (TPF/hr) in the operation.

After these filters have been applied, lagged TPF terms are computed, and the corresponding loss of leading observations noted.[5] A final check is then made to ensure that at least eight observations, which need not be consecutive, remain for each

---

[4] MODS operation groups are referred to in the "dot..." statement by number: '17' = BCS/ (MPBCS), '18' = BCS/DBCS, '19' = FSM/, and '20' = FSM/1000, '01' = OCR, '12' = SPBS, '02' = LSM, '05' = Manual Flats, '06' = Manual Letters, '07' = Manual Parcels, '08' = Manual Priority, '10' = combined BCS, '11' = combined FSM. Within the dot loop, the operation group number is represented by a dot (".") placeholder, which is replaced during loop processing with the appropriate operation group number. See *TSP Reference Manual* (op. cit.) at 70-71. See also Section II, below.

[5] Due to the panel structure of the data, each lagged TPF or TPH term included in the model causes the loss of one leading observation *per facility*. Given that there are initially 321 REGPO positions in the data set, each included lagged term can cause the loss of observations up to a theoretical maximum of 321, or 1,284 in total with all four lagged quarters of TPF in the model.

mail processing facility. If there are fewer than eight observations, the site's data are removed from the regression sample for that operation.

3. Definition of analysis variables and elasticity functions

The functional form used in the estimation process is the transcendental logarithmic ("translog"). See USPS-T-14 at 56-59. Therefore, the analysis variables must be transformed into their natural logarithms (with the exception of the dummy variables and time-trend), and squares and cross-products created. The section of the program following sample selection does the calculations.

Once variable definitions are completed, the program defines the elasticity functions and the various sample aggregation points at which they are to be evaluated. The following elasticities are produced:

Elasticity Name	Elasticity Description
• mvv_.	Elasticity of hours with respect to piece handlings/pieces fed, evaluated at the arithmetic sample mean
• m00vv_.	Elasticity of hours with respect to piece handlings/pieces fed, evaluated at the FY 2000 arithmetic sample mean
• vv_.	Elasticity of hours with respect to piece handlings/pieces fed, evaluated at each sample data point (arithmetic mean of vv_. is equivalent to the elasticity evaluated at the unweighted geometric mean)
• lmvv_.	Elasticity of hours with respect to piece handlings/pieces fed, component from lagged terms only
• lm00vv_.	Elasticity of hours with respect to piece handlings/pieces fed, evaluated at the FY 2000 arithmetic sample mean, component from lagged terms only
• lvv_.	Elasticity of hours with respect to piece handlings/pieces fed, evaluated at each sample data point, component from lagged terms only
Elasticity Name	Elasticity Description
• mdd_.	Elasticity of hours with respect to possible deliveries, evaluated at the arithmetic sample mean

- m00dd_. Elasticity of hours with respect to possible deliveries, evaluated at the FY 2000 arithmetic sample mean
- dd_. Elasticity of hours with respect to possible deliveries, evaluated at each sample data point
- mcc_. Elasticity of hours with respect to facility capital index, evaluated at the arithmetic sample mean
- m00cc_. Elasticity of hours with respect to facility capital index, evaluated at the FY 2000 arithmetic sample mean
- cc_. Elasticity of hours with respect to facility capital index, evaluated at each sample data point
- mww_. Elasticity of hours with respect to LDC wage, evaluated at the arithmetic sample mean
- m00ww_. Elasticity of hours with respect to LDC wage, evaluated at the FY 2000 arithmetic sample mean
- ww_. Elasticity of hours with respect to LDC wage, evaluated at each sample data point
- mpp_. Elasticity of hours with respect to manual ratio, evaluated at the arithmetic sample mean
- m00pp_. Elasticity of hours with respect to manual ratio, evaluated at the FY 2000 arithmetic sample mean
- pp_. Elasticity of hours with respect to manual ratio, evaluated at each sample data point

4. Model estimation, elasticity computation and evaluation

Once the elasticity functions have been defined, the variability model is estimated using the TSP "olsq" command, assuming non-autocorrelated disturbances.[6] The fixed-effects residuals and slope parameters are extracted for further processing. The elasticities and their standard errors under this maintained hypothesis are calculated and evaluated. The Baltagi-Li estimate of the panel autocorrelation coefficient and the

---

[6] In Docket No. R2000-1 I used the TSP "panel" command, which performs a mean-differencing form of fixed effects estimation. In this case, I use the "olsq" command in conjuction with a full set of facility dummy variables. See USPS-T-14 at 40-42 for a discussion.

Bhargava-Franzini-Narendranathan estimate of the panel Durbin-Watson statistic are produced from the residuals.[7] Using the autocorrelation coefficient estimate, and assuming first-order autocorrelated disturbances, the data are transformed so a feasible Generalized Least-Squares (FGLS) version of the model may be estimated. The elasticities and their standard errors from the FGLS model are then calculated and evaluated under this maintained hypothesis. The FGLS elasticities are reported for the automated and mechanized letter and flat sorting operations in USPS-T-14, Tables 9 and 10, at pages 62-63.

B. Manual Letter and Flat Distribution Operations Program

The only material differences between this program, "varltr-tph-by2000.tsp", and the previously-described program for automated and mechanized letter and flat operation groups have to do with the piece handling variable used—TPH rather than TPF (see USPS-T-14 at 58[8])—and the inclusion of the manual ratio and an associated dummy variable that controls for the change in the weight-to-pieces conversion factor (id. at 45-46).

C. Manual Parcel and Manual Priority Operations Program

The only material difference between this program, "varnl-tph-by2000.tsp", and the program for manual letter and flat operation groups described in subpart B is that all terms in the model related to the manual ratio have been dropped. See USPS-T-14 at 58-59.

---

[7] See Docket No. R97-1, USPS-T-14, at 48-50 for the relevant citations.

[8] For a discussion of the differences between TPF and TPH, see Docket No. R2000-1, USPS-T-15 at 50-52.

D. Automated/Mechanized Letter and Flat Operations "F-Test" Program

This program, "varltr-tpf-by2000-ftest", implements the F test of the log-linear ("Cobb-Douglas") model against the more general translog specification for the BCS, FSM, OCR, and LSM distribution operations. See USPS-T-14 at 65-66.

E. Manual Letter and Flat Operations "F-Test" Program

This program, "varltr-tph-by2000-ftest", implements the F test of the log-linear ("Cobb-Douglas") model against the more general translog specification for the Manual Letter and Flat distribution operations. Id.

F. SPBS Operation "F-Test" Program

This program, "varnl-tph-by2000-ftest", implements the F test of the log-linear ("Cobb-Douglas") model against the more general translog specification for the SPBS operation. Id.

G. Manual Parcel and Manual Priority Operations "F-Test" Program

This program, "varltr-tpf-by2000-ftest", implements the F test of the log-linear ("Cobb-Douglas") model against the more general translog specification for the BCS, FSM, OCR, and LSM distribution operations. Id.

H. Automated/Mechanized Letter and Flat Operations Pooled OLS Program

This program, "varmp-tpf-by2000-ols.tsp", implements the pooled OLS regression model for the BCS, FSM, OCR, LSM, and SPBS distribution operations. See USPS-T-14, Appendix A.

I. Manual Letter and Flat Operations Pooled OLS Program

This program, "varltr-tph-by2000-ols.tsp", implements the pooled OLS regression model for the manual letter and flat operations. Id.

J. Manual Parcel and Manual Priority Operations Pooled OLS Program

This program, "varnl-tph-by2000-ols.tsp", implements the pooled OLS regression model for the manual parcel and priority distribution operations. Id.

K. Automated/Mechanized Letter and Flat Operations "Alternative FGLS" Program

This program, "varmp-tpf-by2000-alt.tsp", implements the alternative ("Cochrane-Orcutt") autocorrelation adjustment for the BCS, FSM, OCR, LSM, and SPBS distribution operations. See USPS-T-14, Appendix B.

L. Manual Letter and Flat Operations "Alternative FGLS" Program

This program, "varltr-tph-by2000-alt.tsp", implements the alternative ("Cochrane-Orcutt") autocorrelation adjustment for the manual letter and flat distribution operations. Id.

M. Manual Parcel and Manual Priority Operations "Alternative FGLS" Program

This program, "varnl-tph-by2000-alt.tsp", implements the alternative ("Cochrane-Orcutt") autocorrelation adjustment for the manual parcel and priority distribution operations. Id.

N. Automated/Mechanized Letter and Flat Operations "Alternative Capital Index" Program

This program, "varmp-tpf-by2000-ahe.tsp", substitutes "qiahe" (capital input index for automated handling equipment – i.e., OCRs and BCSs) for "qicap" in the OCR and BCS distribution operations. See USPS-T-14, Appendix D.

O. Automated/Mechanized Letter and Flat Operations Generalized Leontief Program

This program, "varltr-tpf-by2000-gl.tsp", uses the generalized Leontief, rather than the translog, functional form in the OCR, BCS, FSM and LSM distribution operations. See USPS-T-14, at 28 and Appendix C.

P. Automated/Mechanized Letter and Flat Operations "MANR" Program

This program, "varltr-tpf-by2000-man.tsp", includes the manual ratio in estimating equations the OCR, BCS, FSM and LSM distribution operations, to demonstrate the effect of adding the manual ratio variable to the recommended estimating equations. See USPS-T-14 at 47-51.

Q. Manual Letter and Flat Operations "No MANR" Program

This program, "varltr-tph-by2000-noman.tsp", excludes the manual ratio from the manual letter and flat distribution operations, to demonstrate the effect of droppng the manual ratio variable from the recommended estimating equations. Id.

R. Automated/Mechanized Letter and Flat Operations "Panel/Alternative FGLS" Program

This program, "varmp-tpf-by2000-panel", uses the TSP "panel" command, combined with the alternative autocorrelation adjustment described in subsection K, to estimate variabilities for the OCR, BCS, FSM, LSM, and SPBS distribution operations.

The panel command computes the pooled OLS, "between" (group means), fixed effects, and random effects estimators. The F-test statistics reported in USPS-T-14, Table 12, are obtained from the "panel" programs.

S. Manual Letter and Flat Operations "Panel/Alternative FGLS" Program

This program, "varltr-tph-by2000-panel", uses the TSP "panel" command, combined with the alternative autocorrelation adjustment described in subsection K, for the manual letter and flat distribution operations.

T. Manual Parcel and Manual Priority "Panel/Alternative FGLS" Program

This program, "varnl-tph-by2000-panel", uses the TSP "panel" command, combined with the alternative autocorrelation adjustment described in subsection K, for the manual parcel and priority operations.

U. Preliminary Disaggregated Cancellation/Meter Prep Results

This program, "varnl-cancmpp.tsp", produces the preliminary results for the disaggregated cancellation and meter prep operations. See USPS-T-14 at 9-10.

V. "Direct" TPF-FHP Regressions

This program, "tpf-fhp-auto.tsp", reproduces the "direct" tpf-fhp regressions from R2000-1, LR-I-457, modified to include only automated operations. See USPS-T-14 at 30-33.

W. "Reverse" TPF-FHP Regressions

This program, "tpf-fhp-auto-rev.tsp", is the "reverse regression" version of tpf-fhp-auto.tsp". Id.

II. Development of Data Sets for Mail Processing Labor Demand Models

All data processing related to the construction of the analysis data set was performed on a Data General AViiON minicomputer with four Pentium Pro microprocessors and one gigabyte of RAM, running the DGUX version of the UNIX operating system. Source programs ending with a ".f" are FORTRAN programs and programs ending in a ".sm" are SORT/MERGE programs. Programs ending with ".epl" are programs utilizing Christensen Associates' proprietary Economic Programming Language (EPL) software. The remaining processing is performed on PCs running the Windows 2000 operating system and Microsoft Office 2000.

In the definitions and descriptions that follow, <w> is a placeholder denoting week, <ap> is a placeholder denoting accounting periods 1-13, <q> is a placeholder denoting postal quarters (PQ) 1-4, and <y> is a placeholder denoting the two-digit postal fiscal years (PFY) 93-00. The submitted programs are versions run on PFY 00 data. The programs for previous quarters and years are substantially similar. See, e.g., Docket No. R2000-1, USPS-LR-I-107, Section II.

A. Creation of maps used in processing

Maps are needed in this processing to roll data to different levels. Many of the maps are created using the Installation Master File (IMF), which is a file containing all possible Finance numbers and their ZIP Codes. The IMF data are allocated from USPS IBM 3480-compatible cartridge tapes. In earlier years, the data are allocated from 9-track tapes.

Program: **finlist.sm**—Sorts the following list of Finance numbers

Input: **finlist**—List of all Finance numbers with a BA of 1 from PFY 93-00. The only exception is that Customer Service Finance numbers (BA of 4) of Processing and Distribution (P&D) plants are also included. For simplicity, the Finance numbers in this map will be referred to as 'BA 1' Finance numbers.

Output: finlist.s—Sorted list of BA 1 Finance numbers

Program: **finnames.sm** – Creates a text file of Installation Master File (IMF) data

Input: **IMF tapes for APs 03, 06, 09, and 13 for PFY 93-00**
Output: **finnames.<a><y>** - Text file with IMF data (<a> refers to APs 03,06,09, and 13)

Program: **finzip.f** – Uses the IMF to create quarterly maps of Finance numbers to 3-digit ZIP Codes. Finance numbers for area offices and headquarters are not rolled to a 3-digit ZIP Code.

Input: **finlist.s** – Sorted list of BA 1 Finance numbers
districts – List of the 85 Customer Service Districts
ziplist – List of all possible 3-digit ZIP Codes
finnames.<a><y> - Text file with IMF data
Output: **finzip.q<q><y>** - Quarterly maps from all Finance numbers to 3-digit ZIP Codes

Program: **finzipcum.f** – Using the output of the above program, this program creates a cumulative map from PFY 93-00 of all possible Finance numbers to 3-digit ZIP Codes. This program is run for every quarter from PFY 93-00.

Input: **finzip.q<q><y>** - Quarterly maps from all Finance numbers to 3-digit ZIP Codes
finzipcum.q<p><y> - Cumulative Finance number to 3-digit ZIP Code map for previous quarter (<p> = previous quarter)

Output: **finzipcum.q<q><y>** - Cumulative Finance number to 3-digit ZIP Code map. When this program is run for Quarter 4 PFY 00, the name is changed to finzip

Program: **posr** - FOCUS program that creates an annual list of MODS Finance numbers along with their REGPO, FHP, TPH, and hours. This program is run on PFY 93-00.

Input: **USPS Corporate Database MODS File**

Output: **fin<y>** - List of Finance numbers reporting MODS data

Program: **posr.sm**-Sort merge program that creates a list of REGPOs from PFY 93-00.

Input: **fin9300** – Concatenated list of **fin93** – **fin00**

Output: **reg9300** – 421 REGPOs for PFY 93-00

Program: **fin.sm**- Program that sorts the concatenated list of all MODS Finance numbers for PFY 93-00.

Input: **fin9300** – Concatenated list of **fin93** – **fin00**

Output: **posr.map** – Sorted list of MODS Finance numbers from PFY 93-00

B. Preparation of the Address Information System (AIS) Data

The AIS data are allocated from IBM 3480-compatible cartridge tapes created by the United States Postal Service. In earlier years, the data are allocated from 9-track tapes. From AP 01 PFY 93 to AP 06 PFY 94, AIS data are allocated from tapes created by the San Mateo Data Center on an AP basis. Beginning AP 07 PFY 94, tapes created by the Memphis Data Center are used. Quarterly Memphis tapes are used from PQ 3 PFY 94 to PQ 4 PFY 95. Following that period, the Memphis tapes are a bi-monthly frequency. The monthly tape closest to the end of a quarter is used to approximate quarterly variables.

Program: **ais.f** - Rolls up city and rural deliveries by 3-digit ZIP Code. This program also counts total delivery routes for each delivery type

Input: **ziplist** - List of all possible 3-digit ZIP Codes

finzip – Map from Finance number to 3-digit ZIP Code

AP or monthly AIS tape

Output: **aispos.q<q><y>** – Quarterly possible (installed) delivery data by 3-digit ZIP Code

Program: **load.ais.epl** - Reads in all AIS variables by 3-digit ZIP Code and loads all quarterly AIS variables into an EPL data set

Input: **aispos.q<q><y>** - Possible delivery data by 3-digit ZIP Code

Output: **cs.qais<y>z** – EPL data set with all delivery point variables by 3-digit ZIP Code

C. Preparation of the Address List Management System (ALMS) Data

The ALMS data are available monthly on CD-ROM since April of 2000. Previous, ALMS data are allocated from monthly 9-track tapes. The monthly tape closest to the end of a quarter is used to approximate quarterly variables.

Program: **alms.f** - Rolls up the number of Post Offices by type for each 3-digit ZIP Code. This program also counts the number of 5-digit ZIP Codes in each 3-digit ZIP Code.

Input: **ziplist** - List of all possible 3-digit ZIP Codes

finzip – Map from Finance number to 3-digit ZIP Code

Monthly ALMS CD or tape

Output: **pocount.q<q><y>** - Number of Post Offices by type by 3-digit ZIP Code

Program: **load.alms.epl** - Reads in all ALMS variables by 3-digit ZIP Code and loads all quarterly ALMS variables into an EPL data set.

Input: **pocount.q<q><y>** - Quarterly ALMS variables in text file

Output: **cs.qalms<y>z** – EPL data set with all Post Office variables

D. Preparation of the Facility Master System (FMS) Data

The FMS data are allocated from IBM 3480-compatible cartridge tapes created by the United States Postal Service. In earlier years, the data are allocated from 9-track tapes.

Program: **fms.f** - Reads in the FMS Finance number level tape and rolls up USPS owned square footage data by BA 1 Finance number and 3-digit ZIP Code.

Input: **finlist** - List of BA 1 Finance numbers

ziplist - List of all possible 3-digit ZIP Codes

finzip - Map from Finance number to 3-digit ZIP Code

PQ 4 End-of-Year (EOY) FMS tapes

Output: **sqft<y>zip** - 3-digit ZIP Code level USPS owned square footage

sqft<y>fld – BA 1 Finance number level USPS owned square footage

Program: **check.fms.epl** – Prints owned squared footage data across years. This printed data is then checked for errors or omissions.

Input: **finlist** - List of BA 1 Finance numbers

sqft<y>zip - 3-digit ZIP Code level USPS owned square footage

sqft<y>fld – BA 1 Finance number level USPS owned square footage

Program: **load.fms.epl** – Computes quarterly values for owned buildings and land scaled to the NCTB national value by BA 1 Finance number. Annual owned square footage shares from FMS are used to split out the national quarterly values of land and buildings. Corrections to owned square footage are also made.

Input: **finlist** - List of BA 1 Finance numbers

usps.qcap - Data set containing NCTB national values of buildings and land

sqft<y>zip - 3-digit ZIP Code level USPS owned square footage

sqft<y>fld – BA 1 Finance number level USPS owned square footage

Output: **fld.qcapital<y>** - EPL data set containing Finance number level building and land variables

Note: Where obvious omissions or other errors in data entry are visible, a manual correction is included in the corresponding load program. For instance, if the facility at sample one has 20,000 square feet in PFY 93-95 and PFY 97-00, but no square feet in PFY 96, a correction is made giving sample one 20,000 square feet in PFY 96. If that same facility has 40,000 square feet in PFY 96, a similar correction is made.

E. Preparation of the Management Operating Data System (MODS) Data

The MODS data are allocated from the USPS Corporate Data Base. The data are located under the first option at the Corporate Information System (CIS) menu in the database MODSF. This database stores information for roughly 800 Finance numbers which report MODS data.

Program: **oplist** - FOCUS program that creates an annual list of operations. This program is run on each year from PFY 93-00. To get the proper name and function of each operation, this list is downloaded shortly after the year ends. For example, if the PFY 93 operation list were downloaded today, the operations would mistakenly have the same descriptions as the current year.

Input: **USPS Corporate Database MODS File**

Output: **op<y>** - List of Finance numbers reporting MODS data

Program: **modann** - FOCUS program which pulls MODS TPH, TPF, and hours by AP and operation for all reporting Finance numbers. With TPH and TPF it is possible to transfer one quarter's worth of AP data at a time. Hours files are larger and are transferred one AP at a time. Note that **y4** denotes a four digit fiscal year.

Input: **USPS Corporate Data Base MODS File**

Output: **tphq<q><y>** - Finance number level TPH files containing a quarter's worth of AP data later concatenated to create the file **tph<y4>** which contains AP data for one full year

tpfq<q><y> - Finance number level TPF files containing a quarter's worth of AP data later concatenated to create the file **tpf<y4>** which contains AP data for one full year

hrs<ap><y> - AP Finance number level hours files later concatenated to create the file **hrs<y4>** which contains AP data for one full year

Program: **mods.f** - Reads in operation level MODS TPH, TPF, and hours data by Finance number and rolls these data to specified operational groups

Input: **fyap.prn** - List of accounting periods and fiscal years from PFY 93-00

posr.map - List of Finance numbers which report MODS data

opmap<y4> - List of operation numbers and groups the data are rolled to

tph<y4> - Operation and Finance number level MODS TPH

tpf<y4> - Operation and Finance number level MODS TPF

hrs<y4> - Operation and Finance number level MODS HRS

Output: **TPH93_00.fin** - PFY 93-00 MODS operational group level TPH for all reporting Finance numbers

TPF93_00.fin - PFY 93-00 MODS operational group level TPF for all reporting Finance numbers

HRS93_00.fin - PFY 93-00 MODS operational group level HRS for all reporting Finance numbers

The MODS Operation Groups are as follows:[9]

1. OCR
2. LSM
3. SPBS Other
4. SPBS Priority
5. Manual Flats

---

[9] The group numbers are not consecutive because certain operation groups from LR-I-107 are not present in the current data set; their group numbers were not re-used.

6. Manual Letters
7. Manual Parcels
8. Priority
10. BCS
11. FSM
12. SPBS (combined)
13. Cancellations and Metered Prep
17. BCS/ (MPBCS)
18. BCS/DBCS
19. FSM/ (FSM 881)
20. FSM/1000
21. Mechanized Cancellations (MOD 011-015)
22. Hand Cancellations (MOD 010)
23. AFCS (MOD 015)

Program: **load.mods.epl** - Rolls MODS operational group and Finance number level data to REGPO and loads these data to a final data set. MODS data for AMFs, district offices, and VMFs are excluded.

Input: **posrtoreg.map** - List of MODS Finance numbers and their REGPOs
regpos - List of the REGPOs included in the final data set
TPH93_00.fin - PFY 93-00 MODS operational group level TPH
TPF93_00.fin - PFY 93-00 MODS operational group level TPF
HRS93_00.fin - PFY 93-00 MODS operational group level HRS

Output: **regpo.qmods.93_00** – EPL panel data set containing final MODS REGPO and operational group level variables

F. Preparation of the National Consolidated Trial Balance (NCTB) Data

The NCTB data are allocated from IBM 3480-compatible cartridge tapes created by the United States Postal Service. In earlier years, the data are allocated from 9-track

tapes. The following programs are used to extract and create variables for use in the volume-variability regression models.

Program: **nctb.f** - Reads in the NCTB Finance number level tape and rolls up expense data by Finance number and 3-digit ZIP Code.

Input: **nctbmap.s** - Map of NCTB expense accounts

finzip – Map from Finance number to 3-digit ZIP Code

Quarterly NCTB Finance number level tapes

Output: **cfld.q<q><y>** - BA 1 Finance-number-level NCTB revenue account data

Program: **load.nctb.epl** - Reads in Finance-number-level expense data, creates building and equipment rental expense variables, and saves variables to an EPL data set.

Input: **cfld.q<q><y>** - BA 1 Finance number level NCTB revenue account data

Output: **fld.qnctb<y>** – EPL data set containing BA 1 Finance number variables

G. Preparation of the National Workhours Reporting System (NWRS) Data

The NWRS data are allocated from IBM 3480-compatible cartridge tapes created by the United States Postal Service. In earlier years, the data are allocated from 9-track tapes. The following programs are used to extract and create variables for use in the volume-variability regression models.

Program: **nwrs.f** - Reads in the NWRS Finance number level tape and rolls up workhours and compensation data by Finance number.

Input: **finzip** - Map from Finance number to 3-digit ZIP Code

NWRS Finance Number Level Tape

Output: **nwrsfld.q<q><y>** - BA 1 Finance number level NWRS regular and overtime workhours and compensation

Program: **load.nwrs.epl** - Reads in BA 1 Finance number level workhours and compensation and saves all variables to an EPL data set.

Input: **nwrsfld.q<q><y>** - BA 1 Finance number level NWRS regular and overtime workhours and compensation

Output: **fld.qnwrs<y>** – EPL data set containing Finance number level hours and dollars for LDCs 11-14, and 17.

H. Preparation of the Personal Property Asset Master (PPAM) Data

The PPAM data are allocated from IBM 3480-compatible cartridges created by the United States Postal Service. In earlier years, the data are allocated from 9-track tapes. The AP 13 PPAM is used for the following year's processing. For example, the AP 13 PFY 99 tape is used for all quarterly processing in PFY 00. The following programs are used to extract and create variables for use in the volume-variability regression models

Program: **asset.prices.epl** - EPL program that creates a text file with national prices for the four asset types: Customer Service Equipment (CSE), Postal Support Equipment (PSE), Automated Handling Equipment (AHE), and Mechanized Handling Equipment (MHE). The prices are derived from various public and private data sources, as well as USPS sources.

Input: **usps.tfp** – EPL database with national prices for postal equipment

Output: **prices** – Text file with national prices for postal equipment

Program: **ppam.f** - Calculates the distribution of investment from the AP 13 PPAM for all years from 1968 to the current year by asset type. It also calculates the stock for the four asset types: Customer Service Equipment (CSE), Postal Support Equipment (PSE), Automated Handling Equipment (AHE), and Mechanized Handling Equipment (MHE)

Input: **prices** - National prices for postal equipment

pcn.map - List of Property Code Numbers

finzip - Map from Finance number to 3-digit ZIP Code

AP 13 End-of-Year (EOY) PPAM tapes

Output: **zip<y>.stk** - Stocks for MHE, CSE, PSE, and AHE by 3-digit ZIP Code

fld<y>.stk - Stocks for MHE, CSE, PSE, and AHE by BA 1 Finance number

Program: **load.ppam.epl** - Computes quarterly values for CSE, PSE, AHE, and MHE scaled to the national value for each type. Annual shares from PPAM are used to split out the national quarterly values for each asset type.

Input: **usps.qcap** – EPL database containing NCTB national quarterly values for CSE, PSE, AHE, and MHE

zip<y>.stk - Stocks for MHE, CSE, PSE, and AHE by 3-digit ZIP Code

fld<y>.stk - Stocks for MHE, CSE, PSE, and AHE by BA 1 Finance number

Output: **fld.qcapital<y>** - EPL database with quarterly values of CSE, PSE, AHE, and MHE scaled to national values by Finance number

I. Preparation of Remote Encoding Center (REC) Data

The REC data are allocated from Microsoft Excel spreadsheets created by the United States Postal Service. These spreadsheets contain hours and various images by Processing and Distribution (P&D) facility and REC combination. RECs serve more than one P&D but most postal data systems only provide data by the REC Finance number. Therefore, a method is needed to divide REC costs among the plants that it serves. For each REC, a share of hours is calculated for each P&D that the REC serves. These shares are then used to split the REC's costs among the P&Ds that it serves. RECs didn't start reporting data until week 2 of AP 09 PFY 94. REC data for PFY 94 and PFY 95 are processed together.

Program: **reclist.sm** – Sorts the map of P&Ds to RECs by P&D name

Input: **reclistq<q><y>** - Map of P&Ds to RECs

Output: **reclistq<q><y>.s** - Map of P&Ds to RECs sorted by P&D name

Program: **rec.f** - This program reads the weekly REC data. A share of hours is calculated for each P&D that an individual REC serves. These shares are then used to split REC costs among the P&Ds that it serves. REC sites may switch which plants they serve so quarterly maps are needed.

Input: **rectra<l><w>.prn** – Text files created from USPS Excel spreadsheets
<l> = last digit of year
reclistq<q><y>.s – Map of P&Ds to RECs sorted by P&D name
Output: **recdata.q<q><y>** - List of all P&Ds and their percentage of REC hours here <y> in this case means PFY 96-00. PFY 94-95 are processed together to create one map called **recdata.9495** with all of the quarterly shares in both years.

J. Preparation of the Rural Route Master (RRMAS) Data

The RRMAS data are allocated from IBM 3480-compatible cartridge tapes from the United States Postal Service. In earlier years, the data are allocated from 9-track tapes.

Program: **rrmas.f** - Creates a text file with the number of rural regular boxes, centralized boxes, routes, miles on routes, and stops on the routes by 3-digit ZIP Code

Input: **finzip** – Map from Finance number to 3-digit ZIP Code
RRMAS tapes by AP
Output: **boxzipq<q><y>** - RRMAS data by 3-digit ZIP Code

Program: **load.rrmas.epl** - Loads RRMAS data into an EPL data set

Input: **boxzipq<q><y>** - RRMAS data by 3-digit ZIP Code
Output: **cs.qrrmas<y>z** – EPL data set with rural delivery points by 3-digit ZIP Code

K. Creation of final non-MODS databases from which the volume-variability regressions are run

The databases used in the regressions are created using data from the databases documented above. The programs described below collapse these data to REGPO from the BA 1 Finance number or 3-digit ZIP Code level. The final database contains data for 321 REGPOs.

Program: **load.data.epl** - Loads all the data for a REGPO excluding AMFs, VMFs, and District Offices. All data are collapsed to the REGPO level and loaded into a panel data set. Capital variables are loaded into an intermediate data set as they require more manipulation before being loaded into the final panel data set. Data for AMFs, District Offices, and VMFs are excluded from this data set. The owned or rented buildings booked under the Customer Service Finance number of the P&D are excluded.

Input: **cs.qais<y>z** - EPL data set containing AIS delivery data by 3-digit ZIP Code

cs.qalms<y>z - EPL data set containing ALMS post office data by 3-digit ZIP Code

cs.qrrmas<y>z - EPL data set containing RRMAS rural route data by 3-digit ZIP Code

fld.qcapital<y> - EPL data set containing Finance number level capital variables

fld.qnctb<y> - EPL data set containing Finance number level NCTB variables

fld.qnwrs<y> - EPL data set containing Finance number level NWRS variables

rec.q<q><y>(y=97-00) - Map of facilities served by each REC site

rec.96 – Map of facilities served by each REC site for PFY 96

recdata.q<q><y>(y=97-00) - Shares of REC console hours for each facility served by a given REC site

recdata.q<q>96 - Shares of REC console hours for each facility served by a given REC site

recdata.9495 - Shares of REC console hours for each facility served by a given REC site

dzip.map - Maps 3-digit ZIP Code data to BA 1 Finance numbers

fintoreg.map - Maps BA 1 Finance numbers to REGPO level

Output: **regpo.qdata.93_00.cap** – EPL non-panel data set for creating capital index

regpo.qdata.93_00 – EPL panel data set containing final variables

Program: **load.qindex.epl** - Pulls capital variables from the capital data set and uses national prices to create a capital quantity index variable. This quantity index is then loaded into the final panel data set.

Input: **usps.qcap** - EPL data set containing national prices for buildings, land, and equipment

regpo.qdata.93_00.cap - Capital and materials values by REGPO

Output: **regpo.qdata.93_00** - Panel data set containing final variables

III. Development of MODS Productivity Data for Cost Studies

Program: **refmods.f** – Creates a text file of MODS TPH, TPF, and hours by 59 operation groups, 321 REGPOs, and APs for PFY 00.

Input: **posrtoreg.map** - List of MODS Finance numbers and their REGPOs

pools00.prn – Map of operations to 59 operation groups

regpos - List of the 321 REGPOs included in the final data set

tph00 – PFY 00 downloaded operation and Finance number level MODS TPH

tpf00 – PFY 00 downloaded operation and Finance number level MODS TPF

hrs00 – PFY 00 downloaded operation and Finance number level MODS HRS

Output: **TPH00.reg** – PFY 00 MODS TPH by 59 operational groups, REGPOs, and APs

TPF00.reg – PFY 00 MODS TPF by 59 operational groups, REGPOs, and APs

HRS00.reg – PFY 00 MODS hours by 59 operational groups, REGPOs, and APs

The 59 operational groups are as follows:

Group	Description	Group	Description
1	Out ISS	31	FSM881 Key In MMP
2	In ISS	32	FSM881 Key In SCF
3	REC letters keying	33	FSM881 Key In Primary
4	LMLM	34	FSM881 Key In Secondary
5	Out OSS	35	FSM881 OCR/BCR Out Primary
6	In OSS	36	FSM881 OCR/BCR Out Secondary
7	Out BCS Primary	37	FSM881 OCR/BCR In MMP
8	Out BCS Secondary	38	FSM881 OCR/BCR In SCF
9	In BCS MMP	39	FSM881 OCR/BCR In Primary
10	In BCS SCF/Primary	40	FSM881 OCR/BCR In Secondary
11	In BCS Secondary (1 Pass)	41	FSM1000 Key Out Primary
12	In BCS Secondary (2 Pass)	42	FSM1000 Key Out Secondary
13	In BCS Secondary (3 Pass)	43	FSM1000 Key In MMP
14	Manual Out Primary	44	FSM1000 Key In SCF
15	Manual Out Secondary	45	FSM1000 Key In Primary
16	Manual In MMP	46	FSM1000 Key In Secondary
17	Manual In SCF/Primary	47	FSM1000 BCR Out Primary
18	Manual In Secondary	48	FSM1000 BCR Out Secondary
19	Riffle Letters	49	FSM1000 BCR In MMP
20	Flats VCS Keying Out	50	FSM1000 BCR In SCF
21	Flats VCS Keying In	51	FSM1000 BCR In Primary
22	Flats VCS Keying In	52	FSM1000 BCR In Secondary
23	AFSM100 Out Primary	53	Manual Out Primary
24	AFSM100 Out Secondary	54	Manual Out Secondary
25	AFSM100 In MMP	55	Manual In MMP
26	AFSM100 In SCF	56	Manual In SCF
27	AFSM100 In Primary	57	Manual In Primary
28	AFSM100 In Secondary	58	Manual In Secondary
29	FSM881 Key Out Primary	59	Manual In
30	FSM881 Key		

Program: **load.modsreg.epl** – Loads MODS 59 operational group data by REGPO into an EPL dataset. MODS data for AMFs, District Offices, and VMFs are included.

Input: **TPH00.reg** – PFY 00 MODS TPH by 59 operational groups, REGPOs, and APs

TPF00.reg – PFY 00 MODS TPF by 59 operational groups, REGPOs, and APs

HRS00.reg – PFY 00 MODS hours by 59 operational groups, REGPOs, and APs

Output: **regpo.mods_00** - Panel data set containing PFY 00 MODS data by REGPO, AP, and the 59 operational groups

Program: **refrec.f** – Creates a text file of MODS TPH, TPF, and hours by 59 operation groups, 55 REC REGPOs, and APs for PFY 00

Input: **recs.std** – Map of REC Finance numbers to REC REGPOs

pools00.prn – Map of operations to 59 operation groups

recs - List of the REC REGPOs

tph00 – PFY 00 downloaded operation and Finance number level MODS TPH

tpf00 – PFY 00 downloaded operation and Finance number level MODS TPF

hrs00 – PFY 00 downloaded operation and Finance number level MODS HRS

Output: **TPH00.rec** – PFY 00 MODS TPH by 59 groups, REC REGPOs, and APs

TPF00.rec – PFY 00 MODS TPF by 59 groups, REC REGPOs, and APs

HRS00.rec – PFY 00 MODS hours by 59 groups, REC REGPOs, and APs

Program: **load.modsrec.epl** – Loads MODS 59 operational group data by REC REGPO into an EPL dataset.

Input: **TPH00.rec** – PFY 00 MODS TPH by 59 groups, REC REGPOs, and APs

TPF00.rec – PFY 00 MODS TPF by 59 groups, REC REGPOs, and APs

HRS00.rec – PFY 00 MODS hours by 59 groups, REC REGPOs, and APs

Output: **rec.mods_00** – EPL panel data set containing PFY 00 MODS data by REC REGPO, AP, and the 59 operational groups

Program: **yr.scrub.epl** - Calculates productivities (TPF/Hour) by the 59 operation groups, eliminating the top and bottom non-zero 1% of the REGPO/AP productitivities. If TPH is greater than TPF, TPF is set equal to TPH. A text file is created that is pulled into Excel.

Input: **rec.mods_00** - EPL data set containing PFY 00 MODS data by REC REGPO

regpo.mods_00 – EPL data set containing PFY 00 MODS data by REGPO

Output: **yr_scrub.txt** – Text file with productivities for 59 operational groups