An Empirical Model of Labor Demand for Mail Sorting Operations

Prepared for
The Office of the Consumer Advocate
The Postal Rate Commission

Mark J. Roberts
Professor of Economics
The Pennsylvania State University

May 31, 2002

Goal: Measure how labor use responds to changes in output in mail processing plants. Measure the "output variability of labor demand."

1. Develop A Theoretical Model of Labor Demand

A framework to explain the plant's use of manhours of labor

A. Level of labor use to examine - plant, mail shape, sorting operation

B. Definition of output

C. Role of capital

2. Develop an Empirical Model of Labor Demand

Regression equations of the form - $h_{it} = \beta X_{it} + \gamma\, q_{it} + \varepsilon_{it}$

h_{it} logarithm of manhours of labor
q_{it} logarithm of output
X_{it} other control variables - capital

The parameter γ is the output variability that is the focus of estimation

Main Econometric Issue: Is q_{it} correlated with ε_{it} ?

3. Empirical Results - estimates of γ

By Sorting Operation:
Manual letter or flat sorting - not different than 1.0
Mechanized/Automated (FSM, LSM, OCR, BCS) .88 to 1.2

By Mail Shape: Letters .95 to 1.03, Flats .84 to .96

By Plant: .95 to .99

Theoretical Model of Labor Demand in Mail Sorting Operations

Production Function

Plant Outputs: Number of Sorted Letters (L)
Number of Sorted Flats (F)

Each output is produced separately from the others using its own inputs of capital and labor.

Production of Sorted Letters (L)Utilizes:

Labor in Manual sorting (M_L)
Labor in Mechanized Operations (A_L)
Capital (K_L)

The production function for letter sorting is: $L(M_L, A_L, K_L, l)$

The production function for flat sorting is: $F(M_F, A_F, K_F, f)$

Short-Run Labor Demand Functions

Letter sorting:

1) Manhours in manual operations: M_L (WA_L/WM_L , K_L , L)

2) Manhours in mechanized operations: A_L (WA_L /WM_L , K_L , L)

Interpretation: Manhours in each letter-sorting operation depend on the relative wages of the two groups of workers, the capital used in letter sorting, and the total number of letters sorted in the plant.

Important points to notice:

- A demand equation (and output variability) for each letter-sorting operation.

- Capital used in letter sorting (not total plant capital) is the appropriate control.

- Output variable in both demand equations is the number of letters sorted in the plant. It is not specific to a sorting operation. Use the measure of FHP letters in the plant.

- Can aggregate the output variabilities for manual and mechanized operations to get an output variability for letters. Tells us h an increase in the number of letters sorted (L) affects the total use of labor ($M_L + A_L$) in letter sorting.

Flat Sorting:

3) Manhours in manual operations: M_F (WA_F/WM_F , K_F , F)

4) Manhours in mechanized operations: A_F (WA_F /WM_F , K_F , F)

Interpretation: Manhours in each flat-sorting operation depend on the relative wages of the two groups of workers, the capital used in flat sorting, and the total number of flats sorted in the plant. There will be an output variability for each operation that measures the labor demand response to an increase in F.

- Can construct an aggregate output variability for flats.
  How does an increase in F affect the total use of labor ($M_F + A_F$) in flat sorting?

- Can construct an aggregate output variability for the plant.
  How does an increase in F and L affect the total use of labor ($M_L + A_L + M_F + A_F$) in the plant?

Extensions of the model (section III and Table 1 give a complete list):

1. Recognize that there are several mechanized/automated operations.

   Letters - OCR
   LSM
   BCS - further divided into MPBCS and DBCS

   Flats - FSM - further divided into FSM881 and FSM1000

   Implications: Separate labor demand equation for *each* of 8 mechanized operations. Total of 10 labor demand equations. Control for multiple types of capital.

2. Multiple technologies are used at any point in time and mix of technologies changes over time.

   Figures 1 and 2

   Implications: Control for other technologies present in the plant at a point in time. Cannot model the demand for one sorting operation in isolation from the other operations used.

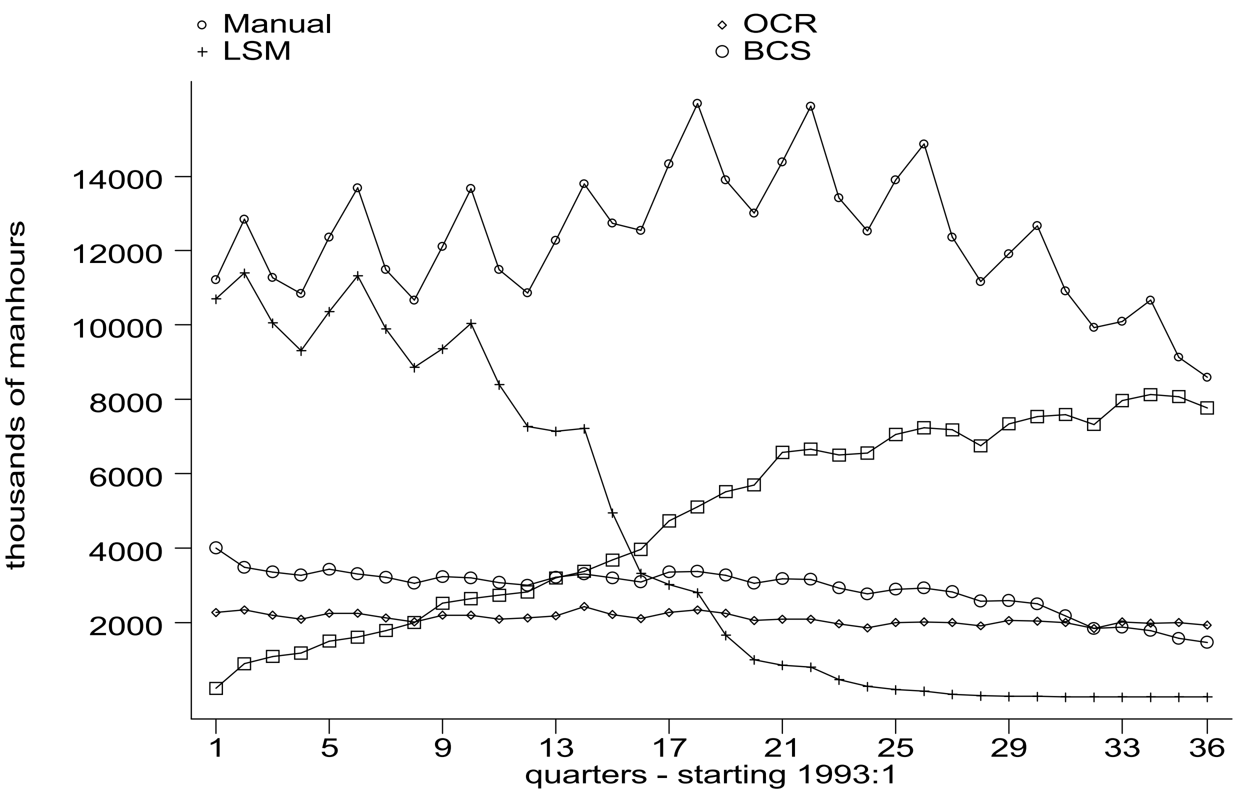
3. There is a change in the measurement of FHP letters and flats in the middle of the time period examined.

   Figures 3 and 4

   Implication: Allow for a shift in the labor demand equations between two regimes.

Figure 1

Total Manhours for Letter Sorting Operations
(Sum over 321 plants)



(The squares represent the DBCS operation)

Figure 2

Total Manhours for Flat Sorting Operations
(Sum over 321 plants)

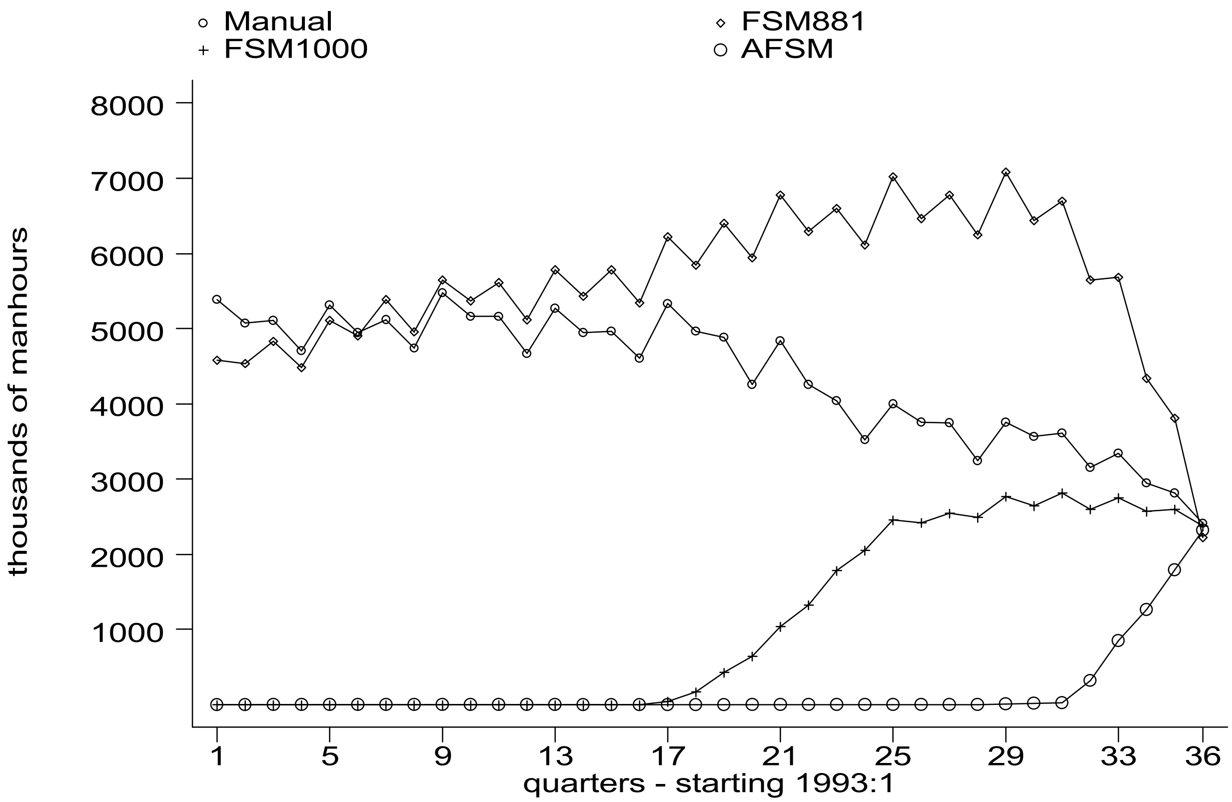


Figure 3

Effect of Change in Conversion Factors on Aggregate FHP Letters
(Sum over 321 plants)

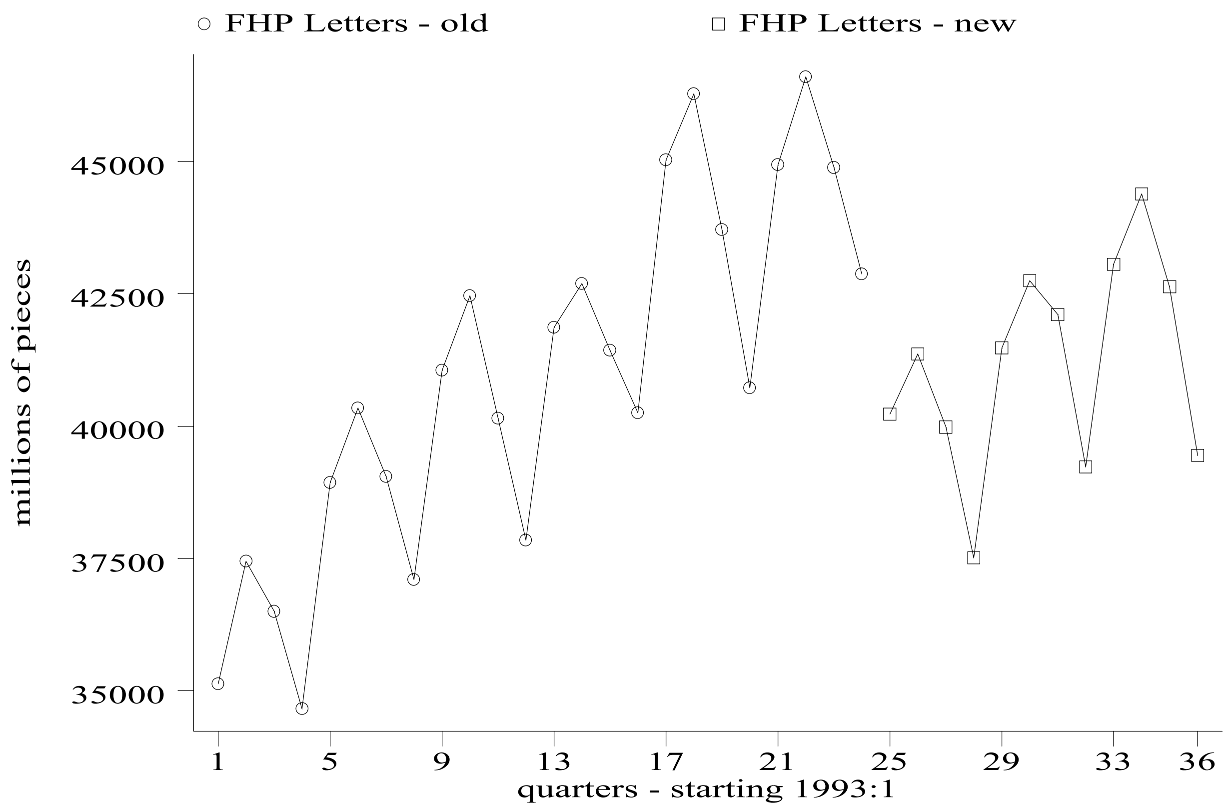
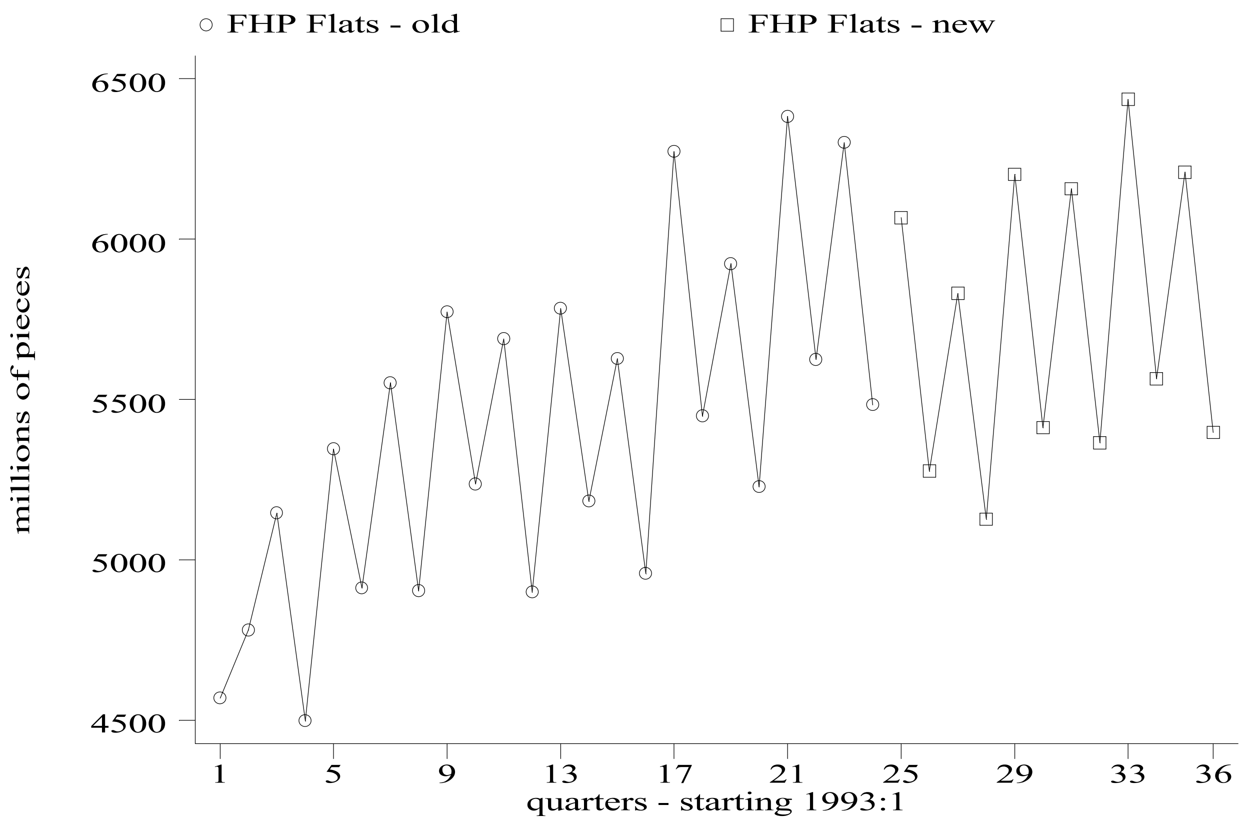


Figure 4

Effect of Change in Conversion Factors on Aggregate FHP Flats

(Sum over 321 plants)



Data and Econometric Issues

Sample of 301 processing plants (out of 321 in data set constructed by USPS in R2001-1). Quarterly observations for 1994-2000. Maximum of 28 time observations/plant.

Basic regression equation (for a sorting operation):

$$h_{it} = \beta X_{it} + \gamma\, q_{it} + \varepsilon_{it} \quad i = 1, 2, \ldots\ldots, 301 \text{ plants}$$
$$t = 1, 2, \ldots\ldots, 28 \text{ time periods}$$

Goal is to estimate the output variability γ for each of the 10 sorting operations.

Key econometric issue: Is $Cov(q, \varepsilon) = 0$?

Yes, output is *exogenous,* ordinary least squares (OLS) estimator has desirable statistical properties.

No, output is *endogenous*, need alternative estimation methods to get desirable statistical properties.

Reasons why q may be endogenous:

1) An important variable m is omitted from the labor demand equation and m is correlated with output.

   Implication: OLS estimator of γ is biased. Direction of bias depends on sign of $Cov(q , m)$.

   Important in this application? Yes.

2) Random measurement error in output.

   Implication: OLS estimator of γ is biased toward zero.

   Important in this application? Yes.

3) Hours and output are chosen simultaneously by plant manager.

   Implication: OLS estimator of γ is biased.

   Important in this application? No.

Correcting for omitted variables and measurement error with panel data

$$h_{it} = \beta X_{it} + \gamma q_{it} + m_i + \varepsilon_{it}$$

1) Remove the effect of time-invariant omitted variables by differencing the data.

Time Differences (First Differences):

$$h_{it} - h_{it-1} = \beta (X_{it} - X_{it-1}) + \gamma (q_{it} - q_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}).$$

Difference from plant means (Fixed Effects):

$$h_{it} - \bar{h}_i = \beta\left(X_{it} - \bar{X}_i\right) + \gamma\left(q_{it} - \bar{q}_i\right) + \left(\varepsilon_{it} - \bar{\varepsilon}_i\right)$$

2) Correct for output measurement error with the use of an instrumental variable (IV).

An IV is used to break the correlation between q and ε.

Variable z can be used as an instrument for the endogenous q if:

- z does not belong in the labor demand equation
- z is correlated with q, after controlling for X
- z is uncorrelated with the regression error ε

Selecting an instrumental variable:

Demand equations for *letter-sorting* operations:

z does not belong in the letter-sorting demand equations (1) or (2)

z is correlated with the true number of letters sorted in the plant L

z is not correlated with random shocks to the hours used in letter-sorting operations.

Good IV's are variables in the *flat-sorting* equations - F and K_F

Demand equations for *flat-sorting* operations.
By similar arguments, good IV's are L and K_L

Econometric methodology

1) Difference the data to remove omitted variables problems. Use both FD and FE.

2) Use instrumental variables estimator to remove measurement error problems. Use F as the instrument for L in letter-sorting operations and L as the instrument for F in flat-sorting operations.

Empirical Results

1) Special case - Manual Flat Sorting Operation in plants with no mechanized flat sorting - Table 3.

OLS estimate is .926.

FE and FD correct for time-invariant omitted variables but can exaggerate measurement error bias.

IV estimates correct for both - estimates are larger than any others and not statistically different than 1.0.

Table 3

Output Variability of Labor Demand for Manual Labor in Flat Sorting
(Estimated using 45 plants that have no mechanized flat sorting)

$$lnH_{05} = \beta_0 + \beta_1 DC + \beta_2 DQ2 + \beta_3 DQ3 + \beta_4 DQ4 + \beta_5 TREND + \gamma \, lnF + \varepsilon$$

Estimator	$\hat{\gamma}$	$SE(\hat{\gamma})$	$\hat{\sigma}^2$
OLS	.926	(.014)	.054
Fixed Effects	.895	(.022)	.015
First Difference	.931	(.023)	.010
Fixed Effects - IV	.981	(.038)	.015
First Difference - IV	.966	(.026)	.010

2) Individual Flat and Letter-Sorting Operations - Table 4.

Comparing estimators:

- difference estimators are always smaller than OLS. Consistent with a downward bias due to measurement error than is exaggerated by differencing.

- IV estimators are always larger than non-IV estimators. Consistent with instruments correcting measurement error bias.

IV estimates of output variability:

Flat-Sorting operations:

Manual .884 to .961, not different than 1.0
FSM .916 to .963, not different than 1.0
Disaggregated FSM operations .348 to .948 , large se for FSM1000

Letter-Sorting operations:

Manual .996 to 1.002, not different than 1.0
LSM, OCR, BCS .882 to 1.218, most not different than 1.0
FD/IV estimates .972 to .992
Disaggregated BCS .682 to 1.241, some different than 1.0

3) Output Variabilities for Letters, Flats, and total Plant - Table 7

Letters .951 to 1.026, not different than 1.0
Flats .838 to .956, some different than 1.0
Plant .952 to .992, one estimate different than 1.0

Table 4

Output Variability of Labor Demand: Alternative Estimators

(robust standard errors in parentheses)

	OLS	Fixed Effects	First Difference	Instrumental Variables	
				Fixed Effects	First Difference
Flat SortingOperations					
Manual (05)	1.365* (.017)	.671* (.072)	.703* (.053)	.884 (.075)	.961 (.044)
FSM - all (11)	.980 (.015)	.615* (.070)	.735* (.063)	.963 (.061)	.916 (.042)
FSM881 (19)	1.086* (.028)	.616* (.053)	.748* (.067)	.803* (.054)	.948 (.040)
FSM1000(20)	.551* (.044)	.381* (.089)	.313* (.071)	.739 (.247)	.348* (.146)
Letter Sorting Operations					
Manual (06)	.853* (.014)	.700* (.049)	.813* (.057)	1.002 (.051)	.996 (.038)
LSM (02)	1.093* (.022)	.654* (.129)	.766 (.131)	1.137 (.188)	.992 (.149)
OCR (01)	1.037* (.018)	.732* (.051)	.815* (.063)	.882 (.084)	.972 (.054)
BCS - all (10)	1.273* (.016)	1.040 (.041)	.879 (.053)	1.218* (.057)	.981 (.043)
BCS (17)	1.610* (.038)	.569* (.096)	.727* (.098)	.682* (.158)	.851 (.094)
DBCS (18)	1.056* (.021)	.837* (.077)	.818* (.066)	1.241 (.161)	.998 (.078)

* Reject that the coefficient equals one at the .01 significance level using a two-tailed test.

Table 7

Aggregate Output Variabilities by Mail Shape and Plant

(Mean and standard deviation of the mean over all plant-time observations)

Estimator	Flats ε_F	Letters ε_L	Total Plant ε
Based on Disaggregated Operations for BCS (17,18) and FSM (19, 20) Categories			
OLS	1.186 (.015) *	1.016 (.009)	1.064 (.008) *
FE	.627 (.042) *	.709 (.034) *	.683 (.027) *
FD	.697 (.039) *	.799 (.037) *	.767 (.029) *
FE-IV	.838 (.046) *	1.026 (.050)	.968 (.038)
FD-IV	.914 (.029) *	.980 (.033)	.958 (.025)
Based on Aggregated Operations for BCS (10) and FSM (11) Categories			
FE-IV	.917 (.049)	1.025 (.031)	.992 (.027)
FD-IV	.956 (.029)	.951 (.023)	.952 (.018)*

* Reject that the coefficient is equal to one at the .01significance level with a two-tailed test.