An Empirical Model of Labor Demand for Mail Sorting Operations

Prepared for

The Office of the Consumer Advocate
The Postal Rate Commission

by Mark J. Roberts

May 31, 2002

Table of Contents

I. Introduction

In R2001-1, USPS-T-14, the Postal Service presents an analysis of the variability of labor demand with respect to changes in the volume of output in mail sorting plants. The estimates are derived from an econometric model of labor demand that is estimated using a panel data set of quarterly observations for the years 1995-2000 for 321 mail sorting plants operated by the USPS. The output variabilities are estimated for eleven different sorting operations within the plant and indicate that the hours of labor employed respond less than proportionately to changes in output. As shown in USPS-T-14 (Table 6, Col.1), the smallest output variabilities are in manual sorting operations, and vary from .44 to .71, while automated and mechanized operations have slightly larger output variabilities that vary from .74 to .94. Aggregating over all eleven sorting operations, they estimate a composite variability for the plant of .71 which indicates substantial increasing returns to labor in mail sorting.

In this paper we develop an alternative model of mail sorting within a Postal Service plant. The theoretical model of production recognizes that different sorting operations may be either substitutes or complements within the plant, that the type of capital equipment used in the plant is an important determinant of the type and amount of labor hours utilized, and implies a definition of output that differs from the measure used by the Postal Service in USPS-T-14. From the production model we derive a set of labor demand equations, one for each sorting operation, that distinguish mail by shape (letters, flats, and parcels). The labor demand equations for all flat and letter sorting operations are estimated econometrically using the same set of mail sorting plants analyzed in the USPS study. We focus on the letter and flat sorting operations because virtually every plant sorts both types of mail and we can construct the necessary output measures for both flats and letters. Unlike the USPS study, we do not analyze parcel-sorting operations in these plants because it is not possible to construct the measure of output we believe is appropriate for parcel sorting. In the final section of the paper, we also undertake a preliminary analysis of the manual and mechanized sorting of priority mail.

The majority of the discussion in this paper will focus on the letter and flat-sorting operations. The empirical results for these two mail shapes indicate that the output variabilities of labor demand are not significantly different than one for manual sorting operations and for several of the mechanized or automated categories. In particular, in the key categories of flat-sorting machinery (mods category 11) and bar code sorters (mods category 10), we do not find evidence of increasing returns to labor. When these operations are disaggregated more finely based on the type of machinery used, we do find evidence of increasing returns to labor in some equipment categories, in particular, flat sorting on the FSM1000 machines (mods category 20), but evidence of constant or decreasing returns on other equipment, specifically, letter sorting on the delivery bar code sorters (mods category 18).

The theoretical model we develop in this paper provides a framework for aggregating across the different sorting operations to produce estimates of the variability of labor demand with respect to an increase in the number of letters sorted in the plant and the number of flats sorted in the plant. Using two different econometric estimators and alternative definitions of the sorting operations, we produce several estimates of the average output variability for letter sorting across all plants. The estimates vary from a low of .951, with a standard error of .023, to a high of 1.026, with a standard error of .050. None of the estimates we construct for letter sorting are statistically different than one . Similarly, aggregating over the flat-sorting operations, our estimates of the average labor demand variability with respect to an increase in the number of flats sorted vary from a low of .838 with a standard error of .046, to a high of

.956, with a standard error of .029. Several of the estimates for flat sorting are significantly different than one, providing evidence of increasing returns to labor, and these occur when the FSM 1000 sorting operation is treated as a distinct category.

Our theoretical framework also provides a way to aggregate over all the letter and flat-sorting operations to construct a composite variability estimate for the entire plant. The plant-level output variability measures the proportional increase in total labor across all flat and letter operations with respect to a proportional increase in the number of letters and flats handled in the plant. Our estimates of this composite variability vary from a low of .952, with a standard error of .018, to a high of .992, with a standard error of .027, depending on the econometric estimator and detailed sorting operations used. Overall, at the plant level we find evidence consistent with constant returns to labor in letter and flat sorting. This finding differs substantially from the composite estimate of .71 constructed in USPS-T-14.[1]

In an attempt to understand the reasons for the difference in results between this study and the USPS analysis, the final section of this paper undertakes a systematic comparison of the differences in model specification between the studies. While there are numerous differences in specification, including differences in the key output, capital, and wage rate variables, the functional form of the regression equation, and the treatment of the substitutability of different sorting operations, the main reason for the difference in results lies in the econometric techniques utilized. The key econometric issue is whether the output variable in the labor demand equation is exogenous, meaning it is uncorrelated with the error term in the regression equation. Using a statistical test, we find that output is not exogenous in the labor demand

[1] This estimate of .71 also includes parcel and priority mail sorting operations which our estimate does not.

equations and we utilize the instrumental variables (IV) estimator to correct for this problem in the data. It is the treatment of this econometric issue that is the primary reason our estimates of the output variabilities are larger than the estimates developed in the USPS study.

The remainder of this paper is organized in the following way. In section II we develop a theoretical model of short-run labor demand for mail sorting operations and discuss several ways in which the demand equations vary from those estimated by the Postal Service in USPS-T-14. In the third section, we discuss how to apply the model to the MODS sorting operations used in postal plants. Specifically, at this point we incorporate information on the mix of technologies used in the plant at a point in time and the type of capital equipment utilized. The fourth section outlines some aspects of the plant-level data that are relevant for the estimation. The fifth section discusses a number of econometric issues, focusing on whether or not the output level is likely to be exogenous in the labor demand equations. Reasons it may not be exogenous include omitted variables, measurement error in output, and simultaneity. The use of plant-specific intercepts to correct for omitted plant variables and the use of instrumental variables estimators to correct for measurement errors and simultaneity are discussed. We illustrate the importance of the econometric issues by analyzing the demand for labor in manual flat sorting for a specialized subset of plants, those that do not use any mechanized equipment to sort flats. This special case allows us to abstract from some of the difficult issues arising from the use of multiple technologies in the plant and identify the econometric issues at work. The sixth section of this paper estimates labor demand equations for ten flat and letter-sorting operations. The seventh section provides a discussion of areas in which the modeling framework can be extended and estimates several alternative models to help isolate the source of differences between the results reported here and the ones presented in USPS-T-14.

II. A General Model of Production and Labor Demand in Mail Sorting Operations

We begin by viewing each mail processing plant as producing three distinct outputs per unit of time: a number of sorted letters (L), sorted flats (F), and sorted parcels (P). The inputs used to produce these outputs are various types of labor, measured as hours in different sorting activities, and capital, represented by the quantity of different types of sorting machinery. A final input into the sorting process is the number of unsorted letters (l), unsorted flats (f), and unsorted parcels (p) that arrive at the plant. These are analogous to the raw materials that are transformed into final output in a manufacturing process. Using its capital and labor, each plant will eventually transform the unsorted inputs l, f, and p into the sorted outputs L, F, and P. Because every piece of mail that enters the plant is sorted, $l=L$, $f=F$, and $p=P$. That is, the number of sorted letters produced is equal to the number of unsorted letters that enter the plant. This distinction is important because it is the outputs L, F, and P that are the conceptually correct measures of output to use in modeling plant production.

The nature of the production process allows us to impose some structure on the model of production. Generally, plants sort the three types of outputs in entirely different ways using different machinery and in different locations in the plant. This implies that a useful simplification is to view the plant as a combination of three production processes, one for each shape of mail. Focusing on the process for sorting letters we can view the plant as producing the output L using a combination of specialized capital equipment (bar code sorting machines, optical character readers, and letter-sorting machines) represented as K_L, manhours running the automated/mechanized machinery A_L , and manhours in manual operations M_L.[2] Notice that A and M both represent manhours, but in different letter-sorting activities. The L subscript denotes that these inputs are all devoted to the sorting of letters. The production function for sorted letters can then be written as: $L(K_L , A_L , M_L , l)$.

The production process for sorting flats and parcels is modeled similarly. The output of sorted flats F is produced with a combination of capital used only in that process (K_F), manhours in automated flat-sorting operations (A_F), and manhours in manual flat sorting (M_F). The output of sorted parcels P is produced with a combination of capital used only in that process (K_P), manhours in automated parcel-sorting operations (A_P), and manhours in manual parcel sorting (M_P). The production process for the whole plant can be represented by the transformation function:

$$T(L(K_L , A_L , M_L , l), F(K_F , A_F, M_F , f), P(K_P , A_P , M_P , p)) = 0. \tag{1}$$

This transformation function embodies the assumption that the production processes for the three outputs are separable, and this will have implications for the form of the labor demand equations.

We will view the plant as making a decision on the amount of labor to allocate to each sorting process and to the manual or automated operations in that process. There are six different types of activities that the workers in a plant can be assigned to, a manual operation and an

[2] At this point we are aggregating all machinery used to sort letters into a single capital input. This is just to simplify the theoretical model. K_L can be viewed as a vector of different types of capital equipment. In the empirical model developed below we will disaggregate the letter-sorting capital into several distinct types. Similarly, there can be multiple types of automated labor or manual labor. All the important distinctions for the empirical model can be made with this simplified framework that recognizes one type of capital, labor in automated/mechanized operations, and labor in manual operations. Again, for simplicity, we will refer to all sorting operations other than manual as automated operations.

automated operation for each of the three mail shapes. That is, the plant chooses the number of manhours A_L , M_L , A_F , M_F , A_P , and M_P to minimize the plant's total expenditure on labor. Each time period when it makes these choices it treats the amount of capital equipment K_L , K_F, and K_P as fixed for that time period. It also treats l, f, and p as the expected quantity of letters, flats, and parcels that will arrive to be sorted in the plant. When the plant chooses these labor inputs it is implicitly deciding how to allocate the sorting of each shape of mail between manual and automated operations.

This description of the plant's technology and choice problem results in a cost function for the plant that is separable in the three outputs. The plant's cost function takes the form:

$$(2) \qquad C(C_L(K_L , L , WA_L , WM_L), C_F (K_F , F, WA_F , WM_F), C_P(K_P , P, WA_P , WM_P)).$$

C is the total expenditure on all labor in the plant, while C_L , C_F , and C_P are the expenditures on labor in letter sorting, flat sorting, and parcel sorting operations, respectively. Each of these shape cost functions depends on the capital input used in that shape, the total amount of sorted output, and the prices of labor used in automated (WA) and manual (WM) sorting of that shape. In this model there are six labor demand functions, one for each of the six types of labor, that can be derived from the plant's cost function. The form of the cost function in equation (2), which in turn results from the assumptions made in specifying the plant's production function (1) and the fixed and variable inputs in the cost minimization problem, implies the relevant arguments for each of the labor demand functions.

In general, the demands for manhours in manual operations and manhours in automated operations to sort a particular shape will be a function of the capital, output, wages paid in manual operations, and wages paid in automated operations *for that shape*. Specifically, the labor demand functions for each of the six types of labor are:

Letter sorting operations

manhours in automated/mechanized operations $A_L (WA_L , WM_L, K_L , L)$

manhours in manual operations $M_L (WA_L , WM_L, K_L , L)$

Flat Sorting operations

manhours in automated/mechanized operations $A_F (WA_F , WM_F, K_F , F)$

manhours in manual operations $M_F (WA_F , WM_F, K_F , F)$

Parcel sorting operations

manhours in automated/mechanized operations $A_P (WA_P , WM_P, K_P , P)$

manhours in manual operations $M_P (WA_P , WM_P, K_P , P)$

This model of production implies labor demands that have a different form than the ones presented in the Postal Service testimony (R2001-1, USPS-T14 and R2000-1, USPS-T15). In the remainder of this section we will identify and discuss the important differences.

II.A The Measure of Plant Output

This model indicates that there are three outputs produced by the plant: the number of unique sorted letters (L), the number of unique sorted flats (F), and the number of unique sorted parcels (P). Each piece of mail is counted only one time since each piece of mail accounts for, at the end of the process, only one piece of sorted output. In practice, the output of sorted letters, flats, and parcels could be measured at the exit door of the plant. Alternatively, because every piece of mail that is received in the plant generates one piece of sorted output, the number of letters, flats, and parcels can also be measured as the number of unsorted pieces received at the entry door of the plant. In the notation of the production model from the last section, we can measure the number of sorted pieces L, F, and P or the number of unsorted pieces, l, f, and p that arrive at the plant. The data set we utilize, which is the one developed by the Postal Service in USPS-T-14, provides measures of l, f, and p, the number of unsorted letters, flats, and parcels that arrive at the sorting plant. These are the measures of first-handling pieces (FHP) that are contained in the data set. Because of the nature of the production process, the counts of FHP letters, flats, and parcels are also measures of the quantities of the three outputs produced by the plant. Variation in FHP letters, flats, and parcels over time and across plants will be the crucial source of output variation that we will utilize to measure the variability of labor demand with respect to output changes.

It is important to emphasize that the measure of output implied by the theoretical model is the same for both the automated and manual labor inputs used to sort a particular shape. In the case of letter sorting, the number of FHP letters in the plant, our measure of the output L, is the determinant of the demand for labor in both manual operations (M_L) and automated operations (A_L). An increase in the plant's output L will lead to a change in both the number of manhours in manual operations (M_L) and automated operations (A_L), and it is this hours response that is the basis for estimating the output variabilities that are the focus of this paper. There is not a separate output measure for the manual operations and the automated operations.

It is also important to emphasize that the plant does not control the volume of mail that arrives to be sorted. Hence the levels of the three inputs l, f, and p and the resulting outputs L, F, and P are all exogenous to the plant.[3] While the volume of FHP letters, flats, and parcels is

---

[3] Whether or not output is exogenous to the plant will be very important when choosing the appropriate econometric method to use in estimating the labor demand equations. This is discussed in section V of this paper.

exogenous to the plant, how to allocate this across automated and manual operations is a choice variable for the plant. The plant manager chooses the mix of labor inputs to use given the volume of letters, flats, and parcels that arrive at the plant to be sorted. Measures of "output" that are *specific to a sorting operation*, such as the number of FHP letters allocated to manual sorting or the number of letters fed into a bar code sorting machine, are not an exogenous measure of the output of an operation. Rather, they are a reflection of the manager's decision on how best to utilize the capital inputs available and how to allocate labor across different sorting operations.

This definition of plant output differs significantly from the output measure used in the Postal Service analysis. They use a measure that is both specific to the sorting operation and, in the case of automated operations, varies with the number of passes through the sorting machinery that the letters undergo. The problem with being specific to the sorting operation is that it is not an exogenous measure of the output produced by the plant, but rather a measure of the manager's input choice, as discussed in the last paragraph. The problem with using machine counts of the total number of pieces fed (TPF) into the sorting machinery is that how many times each piece is fed before it reaches its final sort will depend on the type and capacity of sorting equipment used and the decisions on how to program the sorting schemes. Machinery that can sort to more destinations in a single pass will require fewer feedings of the letters and thus generate a lower count of TPF for a given number of letters. If TPF is used as the measure of output, then it will depend upon the type and capacity of sorting machinery present in the plant.

A simple numerical example may help make the point. Suppose there are 1000 FHP letters arriving at the plant and they need to be sorted to 50 separations. If the sorting machine is large enough to sort to 50 separations then each letter will require a single pass through the machine and there will be an "output" of 1000 TPF. Instead, suppose the sorting machine can only sort to 10 separations at a time. This will require multiple feedings for at least some of the 1000 letters and generate a count of TPF that is greater than 1000. The TPF count is higher because the capacity of the sorting machine is less, not because there are more letters to be sorted. The output of sorted letters is 1000 in either case. This distinction has an important implication for estimates of output variabilities. The operation using the smaller capacity machine will require more manhours since it requires more setups and more handling of each piece. An empirical model that uses TPF as the output measure will attribute at least some of the extra hours as resulting from extra output and this will affect the estimate of the output variability. In reality, the extra manhours reflect the fact that there is less capital to work with. An empirical model that uses FHP as the output measure would not measure any difference in output in the two cases but rather would attribute the differences in manhours to differences in either capital variation or random errors depending on how the rest of the equation is specified.

II.B The role of capital

As seen from the labor demand equations developed above, the capital used to sort letters (K_L) affects the demand for manhours in both automated operations (A_L) and manual sorting (M_L) . In general, capital is a substitute for hours in manual sorting operations but a complement for automated manhours. A plant with a larger capital stock will, other things equal, demand fewer labor hours in manual operations and more hours in automated operations. The demand for labor in manual operations is a function of the plant's input of sorting capital. The model also indicates that it is necessary to use disaggregated measures of capital input. At a minimum the capital input should be disaggregated enough to distinguish capital used to sort each shape. In the labor demand equations in the Postal Service analysis, the total capital stock of the plant is used as the capital input for all letter, flat, and parcel sorting operations. This total is an aggregate of K_L , K_F , and K_P and is not specific to each shape. In USPS T-14 (p. 69 and appendix D) the Postal Service provides estimates for the automated letter-sorting categories (BCS and OCR) using an index of automated letter-sorting capital that is closer to the capital measure K_L in the model developed here.[4] In the model estimated in this paper we will utilize measures of capital input disaggregated by the type of machinery and shape of mail it is used to sort.

II.C The Wage Rate Variable.

In the labor demand model developed above there are two types of labor used to sort each shape. Labor demand functions depend on the *relative price* of the alternative inputs, and the correct explanatory variable in the estimating equations for both automated and manual labor is the ratio of the price of automated to manual labor.[5] As the relative price rises, the cost of using labor in automated operations increases and, if the two types of labor are substitutes, the plant should increase its use of the less expensive manual labor. The Postal Service analysis does not use a relative price of labor in the demand models but rather includes just the wage rate of the labor in the sorting operation under study. This makes it difficult to interpret substitution effects among different categories of labor.

---

[4] They did not find that the capital input, whether total plant capital or the aggregate for BCS and OCR capital, was generally a significant determinant of labor use.

[5] The technical terminology is that input demand functions are homogeneous of degree zero in the prices of the inputs. This means that if all input prices increase by the same proportion, there will be no change in the plant's use of inputs.

II.D The interpretation of output variabilities

The reason for estimating these labor demand curves is to calculate the variability of labor demand with respect to changes in output. Given the framework developed above, it is straightforward to define and interpret the output variability. Each labor demand equation will provide an estimate of the variability of one type of labor hours with respect to a change in the number of pieces of mail of a shape that arrive in the plant. For example, the demand equation for manual letter sorting (M_L) summarizes the proportional change in hours in that operation with respect to a proportional change in FHP letters. This is denoted with the derivative $\left(\frac{\partial \ln M_L}{\partial \ln L}\right)$.

Similarly, the demand equation for manual flat sorting (M_F) will summarize the proportional change in hours in that operation with respect to a proportional change in FHP flats. This is denoted with the derivative $\left(\frac{\partial \ln M_F}{\partial \ln F}\right)$.

For each shape, the relevant sorting operations can be aggregated to provide a measure of the total increase in hours as a result of an increase in the FHP count for that shape. Total hours in letter-sorting operations (H_L) are the sum of hours in manual and automated operations: $H_L = M_L + A_L$. The variability of H_L with respect to a change in the number of FHP letters (L) arriving at the plant is a share-weighted sum of the variability of manual hours and the variability of automated hours:

(3) $$\varepsilon_L \equiv \frac{\partial \ln H_L}{\partial \ln L} = \left(\frac{M_L}{H_L}\right)\left(\frac{\partial \ln M_L}{\partial \ln L}\right) + \left(\frac{A_L}{H_L}\right)\left(\frac{\partial \ln A_L}{\partial \ln L}\right)$$

The shares are the share of total hours in letter sorting devoted to the manual or automated operations. The variabilities of manual and automated operations will be estimated from the labor demand equations and can then be aggregated into a variability of labor-sorting hours using equation (3). It is important to emphasize that ε_L measures the proportional change in total labor used in letter sorting for a proportional change in the number of letters to be sorted. This reflects the total adjustment of manhours in all of the letter-sorting operations in response to an increase in the flow of letters. The output change used to measure both of the component variabilities is the same, a change in FHP letters. In the Postal Service analysis each labor demand equation contains a different measure of output (TPF or TPH in the specific operation) and thus each component variability is not measured with respect to the same output. They cannot be aggregated using equation (3) to construct an overall measure of the variability of letter-sorting labor with respect to an increase in the number of letters sorted.

The variabilities of labor demand for hours in flat-sorting operations and in parcel sorting can similarly be constructed as an hours-weighted sum of the variabilities of the manual and automated operations for each shape:

(4) $$\varepsilon_F = \frac{\partial \ln H_F}{\partial \ln F} = \left(\frac{M_F}{H_F}\right)\left(\frac{\partial \ln M_F}{\partial \ln F}\right) + \left(\frac{A_F}{H_F}\right)\left(\frac{\partial \ln A_F}{\partial \ln F}\right)$$

$$\varepsilon_P = \frac{\partial \ln H_P}{\partial \ln P} = \left(\frac{M_P}{H_P}\right)\left(\frac{\partial \ln M_P}{\partial \ln P}\right) + \left(\frac{A_P}{H_P}\right)\left(\frac{\partial \ln A_P}{\partial \ln P}\right)$$

Each of the three shape variabilities ε_L, ε_F and ε_P measures the responsiveness of labor hours with respect to a change in *one* of the plant's three outputs, letters, flats, or parcels, holding the other two outputs constant.

These shape variabilities can be aggregated into an overall labor variability for the plant. This is a way of measuring the total response of labor when there is a proportional change in *all three* of the plant's outputs. The total hours of labor used in all three sorting operations in the plant (H) is the sum of the hours in the three separate operations: $H = H_L + H_F + H_P$. The output variability for total hours is a share-weighted sum of the three shape variabilities:

(5) $$\varepsilon = \left(\frac{H_L}{H}\right)\varepsilon_L + \left(\frac{H_F}{H}\right)\varepsilon_F + \left(\frac{H_P}{H}\right)\varepsilon_P$$

The variability for each of the three shapes as well as the plant's mix of manhours in letter, flat, and parcel sorting operations will contribute to the overall plant variability. Equation (5) provides a useful measure of the total response of the plant's manhours in all sorting operations as the flow of letters, parcels, and flats increases.

Overall, the labor demand framework developed in this section provides a clear way of defining what is meant by a change in the plant's output and a clear way of defining and aggregating the output variabilities for the different sorting operations in the plant. After we estimate the output variabilities for each sorting operation, we will aggregate them to construct a variability for each shape using equations (3) and (4) and for each plant using equation (5).

III. Labor Demand Equations for MODS operations

The production model developed in the last section recognizes some of the key aspects of mail sorting technology, specifically, the separability of the mail stream into different shapes and the use of both manual and automated operations in the sorting of each shape. The model was stylized in restricting attention to just a single type of automated operation for each shape with a single type of capital used. In practice, there are several technologies that are used in the automated sorting of letters, flats, and parcels. There is no difficulty in extending the production model to incorporate several different automated operations. There will be a separate labor demand equation for manhours in each of the operations. The capital stock in each of the automated operations will enter as an additional explanatory variable in all of the labor demand equations for a particular shape. There will be an additional relative wage variable, measuring the wage in each automated operation relative to the wage in manual operations, that will enter into all of the labor demand equations for a particular shape. In this section of the paper we will examine the letter-sorting and flat-sorting operations in more detail and show how the general model from the last section can be tailored to some of the unique features of these sorting operations.

III.A Letter Sorting Operations

The MODS data used by the Postal Service in their analysis disaggregates the sorting of a particular shape into several operations, and we rely on this disaggregation to specify the labor demand equations we estimate. In letter sorting, there are manual operations (MODS category 06), and four different automated operations, optical character readers (01), letter sorting machinery (02), bar code sorters (17), and delivery bar code sorters (18). There is a total of five disaggregated labor demand equations for letter sorting. In addition, the USPS study also aggregates together categories 17 and 18 to create a category for all bar code sorters (MODS category 10). We will also estimate a labor demand equation for this category, giving a total of six labor demands for letter sorting.

The explanatory variables in these labor demand equations will all be identical. The first is the count of FHP letters in the plant (L). This is measured as the sum of FHP letters over all five of these operations. There are five different categories of capital equipment that we can identify in the plant. First, is letter sorting machinery (LSM) and we will measure a capital stock variable for this input. Each plant also reports capital expenditures for bar code readers (BCR) and optical character readers (OCR) and we will construct capital stocks for each of these categories. Ideally, we would like to include separate capital stocks for bar code sorters (BCS corresponding to MODS category 17) and delivery bar code sorters (DBCS, MODS category 18), but the capital data provided does not distinguish these two types of capital and we are forced to specify a single BCS capital stock variable that is the sum of BCS and DBCS equipment. Finally, there is capital equipment that is not specific to the letter-sorting operations. This includes conveyor systems, platform equipment, scales, fork lifts, and other equipment that is not specific to an operation in the plant. We will include an additional capital stock variable to measure this input. Overall the five capital stock variables included in the five letter-sorting demands are: optical character reader capital (KOCR), bar code reader capital (KBCR), bar code sorter capital (KBCS), letter sorting machinery capital (KLSM), and all other capital (KOTHER).

There are two relative wage variables that are included in the demand equations. Following the Postal Service study (USPS-T-14, table1, p.35), we are able to specify the average hourly wage for workers in automated categories 17 (BCS), 18 (DBCS), and 01 (OCR), a separate wage for workers operating letter-sorting machinery (02), and a third wage for manual operations (06). We divide each of the first two categories by the wage in manual operations to create two relative wage variables (WAUT/WMAN) and (WLSM/WMAN). An increase in either relative wage thus represents an increase in the hourly cost of labor for automated or mechanized operations relative to manual operations. This should have a positive effect on the hours used in manual sorting and a negative effect on the level of hours in the automated labor categories.

III.B Flat Sorting Categories

There are three MODS categories that are relevant for the sorting of flats: manual operations (05) and two automated operations, the model 1000 flat sorting machinery (MODS category 20) and other flat sorting machinery (19). This latter include the 881 flat sorting machinery and we will refer to this category as FSM881. In additions, the USPS analysis aggregated categories 19 and 20 together to create a single category for all flat sorting machinery (MODS operation 11). We will also estimate a labor demand equation for this category, giving a total of four labor demand equations.[6]

The explanatory variables include the level of output (F), measured as the sum of FHP flats for the three disaggegated MODS operations, a set of capital stock variables, and the ratio of the average hourly wage of workers operating the flat sorting machinery to the wage of workers in manual operations (WFSM/WMAN). We will include three capital stock variables. The first measures capital stocks in flat-sorting machinery (KFSM). Ideally, we would like to include

---

[6] Beginning in 2001, automated flat sorting machinery (AFSM) was introduced as an additional automated technology. In this analysis we will end in 2000 and not model the use of this new automated technology. In 2001 the adoption of this technology had a major impact on the mix of labor hours used (see Figure 2) and it will be necessary to incorporate this into labor demand estimates that go beyond 2000.

separate measures of the capital stock in the 881 sorting machinery and the 1000 sorting machinery, but the capital data provided by USPS does not allow us to make this distinction and we have to aggregate them into a single capital stock measure. We will also include the capital stock of bar code readers (KBCR) because the reported capital in this category aggregates equipment used in either letter or flat sorting operations (see Response of USPS to interrogatories by UPS in UPS/USPS - T39-61). The final capital stock will be the measure of all other capital used in the plant (KOTHER) because this can be utilized in both letter and flat sorting operations. Overall, there are eight labor demand equations for letter and flat-sorting operations that will be estimated in this analysis.[7]

III.C Multiple Technologies

The model developed in section II of this paper is also stylized in that it describes a single plant that uses all the manual and automated operations. In practice, the mix of technologies is not the same across all time periods or across all plants at a point in time and this will require some additional adjustments in the specification of the labor demand equations. Over time, the postal plants have shifted among different technologies. Particularly important in the time period we will examine (1994-2000) is the phase out of letter sorting machinery and the growth in the use of delivery bar code sorters. Thus there is a reduction in the demand for labor in MODS category 02

---

[7] There are three additional sorting operations in the plant that were included in the Postal Service analysis. They are manual parcels (07), manual priority (08), and the small parcel bundle sorter (12). To include these three operations we need to measure the FHP count of parcels and priority pieces received in each plant. From the data provided in the Postal Service analysis it is not possible to measure the FHP count of parcels in the plant. See the response of USPS to interrogatories of UPS in UPS/USPS-T-39-38 to 41. We will not be able to estimate a comparable model for parcel operations. The Postal Service was able to include the parcel operations in their analysis because they defined output as TPF in each sorting operation which is included in the data set. It is possible to measure the FHP for priority mail. In section VII.E we will provide preliminary estimates for two priority mail sorting operations: manual (08) and priority mail sorted on the SPBS (04).

(LSM) and an increase in category 18 (DBCS) that is related to this shift in technology. Also important in our time period is the deployment of the FSM 1000 machinery that has led to growth in the number of hours in MODS category 20 and a reduction in manual hours, category 05, for sorting flats.

Figure 1 illustrates the importance of these technology changes on the number of hours employed in different letter-sorting operations. It graphs the total hours used in the 321 plants by quarter for each of the five MODS labor pools. The horizontal axis runs from the first quarter of 1993 (time=1) until the fourth quarter of 2001 (time=36). The reduction in the use of letter sorting machinery beginning in 1995 and continuing through 1998 is very clear. Similarly, the increased reliance on DBCS operations over the entire period is clear (the DBCS series is unlabeled on the figure but denoted by the squares). The technology was just beginning to be adopted at the start of the data in 1993 but grew until it was approximately equal to the largest employment category (manual) by the end of the nine-year period. The remaining two automated categories, OCR(01) and BCS (17), show declines over the entire period. Manhours in OCR operations declined by 16 percent over the nine-year period. Manhours in BCS operations declined more dramatically, by 74 percent, with most of the decline in the last two years of the time period. Finally, manual hours began to decline in 1999. Overall, there is a clear initial shift away from letter sorting machinery toward manual and DBCS sorting, followed by a further shift from manual and BCS toward DBCS operations.

The aggregate patterns in Figure 1 reflect both the diffusion of the technology across plants as well as the intensity with which it is used in each plant. Plants do not adopt the technology at the same time. Table 2 summarizes the pattern of technology use for the 321 sorting plants. It reports the proportion of plants that use each technology in the first quarter of each year.[8] The decline in the use of LSM technology is obvious beginning in 1997 and accelerating in 1998. There is a gradual decline in the use of OCR and BCS technologies. The decline in the number of plants using the BCS technology is less dramatic than the decline in hours for that technology, indicating that plants maintain a capability to use the technology (unlike LSM) but use it less heavily over time. Finally, the increase in the use of DBCS technology indicates that it continued to be adopted by plants from 1993 though 1997. The implication of this for the specification of labor demand equations is that to understand the plant's demand for any one category of labor it will be necessary to control for the mix of technologies used in the plant at that point in time. For example, in 1994 letters in the plant would be sorted with manual labor and manhours on letter sorting machines. That same level of output in 1999 would be sorted with a different amount of manual labor and manhours on DBCS machines. The demand for manual labor is likely to be different in the two periods because of the different technologies employed, even if the plant's output was the same in each period. Similarly, the demand for manhours to operate BCS machinery will depend on whether the plant uses the DBCS technology.

The same pattern of technology diffusion is evident in the way plants sort flats. Figure 2 reports the total manhours in each technology. Through 1996 (time=16), flats were sorted with manual labor and the FSM 881 technology. Beginning in 1997 the FSM1000 technology was introduced as a substitute for manual sorting, and this substitution is clearly evident in the time series pattern in Figure 2. The diffusion of this technology across plants was not immediate. Table 2 shows that 4.4 percent of the plants were using the FSM1000 technology at the beginning

[8] The plant is defined as using the technology in a given time period if it reports a positive value for total pieces fed (TPF) if it is an automated operation and total pieces handled (TPH) if it is a manual operation.

of 1997, 38.0 percent at the beginning of 1998, and 60.7 percent at the beginning of 1999.[9] The implication for models of labor demand is that the manhours used in any one category will depend on what other technologies are in use in the plant.

One general way to control for consistent changes in technology over time is to include a time trend among the explanatory variables. Because the dependent variable is measured in logs, the coefficient on the time trend is interpreted as the (constant) per-period growth in manhours after controlling for all the other observable characteristics. It is generally interpreted as the effect of gradual changes in technology on labor demand. It can be either positive or negative, indicating a systematic increase or decrease in the number of manhours as the technology evolves over time.

However, because plants adopt or abandon the technologies at different points in time, it is not sufficient to control for these changes by just including a time trend in the empirical estimating equations. Rather it is also necessary to include some plant-level variables in the model to control for the mix of technologies present in the plant at each point in time. We assume that, within a time-period of observation (a quarter), the mix of technologies present in the plant is fixed.[10] To control for the technologies available to the manager we define a set of dummy variables, one for each major technology, that take the value zero if the technology is not used in the plant and the value 1 if it is. In the case of letter sorting, we define three technology dummies: TECH02

---

[9] More recently, the AFSM technology was introduced beginning in the second quarter of 2000. Compared with the other technologies, its diffusion across plants was fairly rapid. It had been adopted by 31.8 percent of plants by the beginning of 2001 and 47.0 percent of plants by the fourth quarter of 2001.

[10] This does not imply that the mix of technologies is fixed over time or that the manager's allocation of labor across sorting operations cannot change over time in response to changes in the set of technologies available. In addition, for some data reasons discussed in section IV.C, we will not use observations for the time periods in which the plant is introducing or eliminating one of the major technologies. Thus it is reasonable to specify the mix of technologies present in the plant during a quarter as exogenous in the labor demand model.

indicates if the plant uses letter sorting machinery, TECH17 indicates if it uses bar code sorters, and TECH18 indicates if it has adopted the DBCS technology. All of the dummies are based on whether or not the plant reports a positive value for TPF for that technology, and therefore is actually using the technology to sort letters. In the case of flats, we define two technology dummy variables. TECH19 takes the value 1 if the plant uses the FSM881 technology and TECH20 takes the value 1 if the plant uses the FSM1000 technology. Again, the dummies are defined based on whether or not the plant reports a positive value for TPF for that technology. These technology variables capture only whether or not the technology is present and being used in the plant and not how much of the plant's sorting it accounts for or how many hours of labor are devoted to it. The plant's capital stocks for each type of technology will capture information on how heavily each plant is invested in a technology.

III.D Cyclical Fluctuations in Manhours

From Figures 1 and 2 a second pattern is evident in the aggregated hours data. There are strong cyclical patterns in the number of manhours used. For letter sorting this pattern is most obvious for hours in manual operations, where there is peak in the second quarter of each year and a trough in the fourth quarter. It is also obvious for hours in the LSM category in the first three years of the data when that was the dominant automated technology. Over the last four years of the sample, the cyclical pattern is also evident in the manhours in the DBCS category. In flat sorting operations, the cyclical pattern is strong in manual labor throughout the period, in the FSM881 labor in the early years of the sample, and in the FSM1000 labor later in the sample. In flat sorting, the cyclical peak in the use of hours occurs in the first and third quarters, with the troughs in the second and fourth.

The cyclical movements in total hours are correlated with cyclical fluctuations in the volume of mail. On average from the first to second quarter of the year, total FHP letter volume grows 3.3 percent. Over the remainder of the year the average quarterly growth rates are -3.2, -6.2, and 8.2 percent. For flats, the quarterly variation is more substantial. Beginning with the growth rate from the first to second quarter of the year, the four average quarterly growth rates in FHP flats are -10.4, 10.0, -13.4, and 17.6 percent. The cyclical fluctuations in hours can, however, reflect more than variation in output. Differences in work effort or changes in the mix of more- and less-skilled workers may also occur from quarter-to-quarter. In order to control for these other potential sources of variation in hours we will include a set of three quarterly dummy variables (DQ2, DQ3, and DQ4) to identify observations in the second, third, and fourth postal quarters, respectively.

III.E Change in the Measurement of FHP Output

The output measures, FHP letters and flats, are constructed by multiplying the measured weight of letters and flats received in the plant by national conversion factors (pieces/lb) that convert them to piece counts. As discussed in the Postal Service analysis, (USPS-T-14, p.43), the conversion factors were changed at the beginning of 1999. This has the harmful effect of introducing a change in the definition of output between the fourth quarter of 1998 and the first quarter of 1999. The effect of this change can be seen in the plot of the aggregate level of FHP letters in Figure 3. The figure plots the sum of FHP letters over all 321 sorting plants against time. The adoption of the new conversion factors in the first quarter of 1999 (time=25) results in a decline in measured FHP letters. To summarize the magnitude of the change, we regressed the log of FHP letters on a time trend, quarterly dummy variables, and a dummy variable for the time periods after the conversion change. The coefficient on the conversion dummy variable is -.179, which can be interpreted to mean that, after controlling for the upward trend in letter volume over time and the systematic quarterly fluctuations, the change in conversion factors amounted to a 17.9 percent decline in measured output.[11]

The change in the conversion factors also affects the measurement of FHP flats. The quarterly total for FHP flats is graphed in figure 4. While there is a noticeable decline in measured output starting in the first quarter of 1999 (time=25), it is not as substantial as the decline in letters observed in figure 3. Controlling for a time trend and quarterly dummies, the summary regression of log FHP flats produces a coefficient on the dummy variable for the time periods following the definition change of -.110, an 11.0 percent decline in measured FHP output.

There are separate conversion factors for different categories of mail such as machine canceled, metered, or hand canceled. The FHP count we observe for a particular sorting operation in a plant will be the product of the measured pounds of mail in a category and the conversion factor for that category summed over all the categories. Changing the conversion factors for any or all of the mail categories simply causes a multiplicative scaling of the original FHP measure, for a given bundle of mail. Since the regressions will use the logarithm of FHP as the explanatory variable, this is equivalent to adding a constant to the explanatory variable. The average change in the log of FHP resulting from the change in conversion factors can be controlled for in a linear

[11] This coefficient summarizes all factors that make the quarterly observations for 1999 and 2000 different from earlier quarters (after controlling for a time trend and quarterly variation). The change in conversion rates may not be the only factor contributing to this 17.9 percent decline although it is certainly a major reason.

regression by allowing the *intercept* of the regression to differ in the two regimes.[12] This is accomplished by including a dummy variable which is defined as DC=1 if year=1999 or 2000 and 0 otherwise. Unlike the Postal Service analysis, we will be estimating some of the labor demand models as growth rates or first- differences over time. The shift in definition only affects the one growth rate between the fourth quarter of 1998 and the first quarter of 1999. When estimating growth rate models we will simply add a dummy variable (CDC) to distinguish the observations corresponding to the 98:4 - 99:1 difference.

III.F Number of Delivery Points

In the labor demand model estimated by the Postal Service the number of delivery points in the plant's service area is included as an additional explanatory variable. The number of delivery points is defined in USPS- LR-J-56, p.6. It is included to recognize that the geographic area served by each plant will differ. The theoretical model developed above does not recognize any explicit role for characteristics of the service area to affect labor demand. In industries that deliver output over a geographic area, such as electricity, telephone, cable television, and water utilities, the characteristics of the service area, particularly the number of customers and their density, will affect how the firm delivers its output. These characteristics will affect the firm's demand for

[12] The Postal Service study allows all slope coefficients related to output to differ between the two regimes but does not allow the intercept to differ (see USPS-T-14, eq. 16).

inputs.[13] It would certainly be important to account for service area characteristics in modeling local mail delivery. For example, a large thinly-populated geographic area will likely require more delivery manhours per letter delivered than a densely-populated urban area. However, it is less clear that service area characteristics should play a role in determining labor demand within the sorting plant. One way they could matter is if the degree of mail sorting differs across plants depending on how many delivery points they serve. If a plant serving many delivery points has to sort the mail to more final destinations, then they will require more inputs for a given number of letters sorted. In effect, the quality of the output, measured by the depth of sorting, varies depending on the number of destinations served by the plant, and this would affect labor demands. In this case the number of delivery points will act as a proxy for the depth of sorting and could be important in labor demand equations. In contrast, if how the plant sorts mail is not affected by what happens to the incoming mail once it leaves the sorting plant, there is no reason to include service area characteristics in the labor demand equations. Without more information on the depth of sorting operations in the plant, it is not possible to conclude whether the number of delivery points should be included in the labor demand equations. We will initially report estimates without the delivery point variable but will undertake some sensitivity tests in section VII.B to determine whether or not this specification issue is important.

III.G Use of the manual ratio

[13] Roberts, Mark J. "Economies of Density and Size in the Production and Delivery of Electric Power," *Land Economics*, Vol. 62, No. 4, November 1986 estimates a production model for electricity generation and delivery that allows the size of the geographic area and number of customers served by the utility to affect costs and input demands. He also defines measures of economies of scale, density, and size that recognize how expansions in output, number of customers and size of the service area affect costs. Caves, Douglas W., Laurits R. Christensen, and Michael W. Tretheway, "Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ," *Rand Journal of Economics,* Vol. 15, Winter 1984 examine similar issues in the airline industry.

The demand models for manual categories estimated in the Postal Service analysis include an explanatory variable that measures the fraction of the plants' total piece handlings that are done in manual operations. This variable is labeled the "manual ratio" and is included in the empirical equations to control for the overall importance of manual sorting operations in the plant. Since this variable is based on the distribution of pieces across operations, it reflects the plant manager's choice of how to allocate the incoming mail stream between manual and automated operations. It is not an exogenous variable that should be included in the labor demand model. The production model developed in section II of this paper treats manual hours (M) and hours in automated operations (A) for a particular shape as simultaneously chosen, with the mix depending on the plant's capital stock, the relative wage of the two types of labor, and the total number of pieces of mail to be sorted. There is no role for a variable like the manual ratio in the labor demand equations and we will not include it in our specifications.

III.H Final Estimating Equations

Denoting H_J as the manhours in MODS operation *J*, the final labor demand equations, written in implicit form, are:

Letter Sorting Operations

OCR: $H_{01} = H_{01}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH02, TECH17,TECH18)

LSM: $H_{02} = H_{02}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH17, TECH18)

Manual: $H_{06} = H_{06}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH02, TECH17, TECH18)

BCS-All: $H_{10} = H_{10}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH02, TECH17, TECH18)

BCS: $H_{17} = H_{17}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH02, TECH18)

DBCS: $H_{18} = H_{18}$ (L, KLSM, KOCR, KBCR, KBCS, KOTHER, WAUT/WMAN, WLSM/WMAN, TECH02, TECH17)

Flat Sorting Operations

Manual: $H_{05} = H_{05}$ (F, KFSM, KBCR, KOTHER, WFSM/WMAN, TECH19, TECH20)

FSM-All: $H_{11} = H_{05}$ (F, KFSM, KBCR, KOTHER, WFSM/WMAN, TECH19, TECH20)

FSM881 $H_{19} = H_{19}$ (F, KFSM, KBCR, KOTHER, WFSM/WMAN, TECH20)

FSM1000 $H_{20} = H_{20}$ (F, KFSM, KBCR, KOTHER, WFSM/WMAN, TECH19)

Notice that, when specifying the labor demand for hours in category *J*, we do not include the technology dummy variable for technology *J*. This is because the demand equation will be estimated using the plant-quarter observations where the hours used in category *J* are positive. By construction all of these plants will be using technology *J*. The purpose of the technology dummies is to control for the other technologies used in the plant at that point in time. In addition, as in the models estimated by the Postal Service, all estimating equations will include a time trend to control for common shifts in technology experienced by all plants and quarterly dummy variables to control for the systematic seasonal shifts in mail volume. Finally, the hours and output variables will be expressed as natural logs. The capital variables will not be expressed in logs because there are many situations where the plant does not use a particular type of capital and we do not want to drop these observations from the sample.

The differences in model specification between the framework developed in this paper and the one employed in the Postal Service analysis (R2001-1, USPS-T-14) are summarized in Table 1. Each row of the table summarizes how one of the key variables is treated in the two models. The final row of the table summarizes differences in econometric methodology that will be developed below in Section V.

IV. Data Issues

The data used in this analysis has been provided by the Postal Service in USPS-LR-J-56. It has been updated with the revisions provided in LR-J-179 and augmented with the quarterly capital data provided in LR-J-190. It refers to observations for postal quarters over the period 1993-2001 for 321 mail sorting plants. The demand equations will be estimated using the data for 1994-2000, which are the only years where capital data are available.

IV.A Output, Manhours and Wages

The output variable used in the letter-sorting demand equations is the sum of the first-handled pieces (FHP) reported in the operations 01, 02, 17, 18, and 06. The output variable used in the flat-sorting operations is the sum of the first-handled pieces (FHP) reported in the operations 05, 19, 20, and AFSM.

We utilize the same manhours and wage variables as the Postal Service in USPS-T-14. There are three different wage rates that can be constructed in each plant, a wage for manual operations, one for operations with mechanical letter sorting or flat sorting machinery, and one for operations with automated OCR or BCS machinery.[14] As mentioned above, we enter the wage rates in the labor demand equations as the relative wage of the latter two categories to the manual wage.

IV.B Capital Stocks

To construct capital stock measures for the various types of capital inputs used in the plant we begin with data on the plant's expenditures on each piece of capital equipment categorized by a six-digit property code (PCN). The data sets list, for the equipment in operation in a quarter, the original purchase price of the equipment and the year of acquisition. For each plant, time period, and acquisition year, we aggregate these expenditures to seven categories using the 3-digit level PCN codes. The 3-digit PCN categories are Letter Sorting Machinery (910), Flat Sorting Machinery (920), Parcel Sorting Machinery (930), Bar Code Sorters (950 except for category 950000), Bar Code Readers (950000), Optical Character Readers (960), and Other Capital

[14] This is one area in which the data set contains numerous missing observations. Of the maximum 8428 plant-quarter observations we could use, the wage for operators of mechanical letter or flat sorting equipment is missing 933 times. Since this wage rate is used in all the labor demand equations it leads to the loss of approximately 11 percent of the observations. Development of alternative data sources for the wage rates would be helpful in recovering many of these observations.

Equipment (other 3-digit PCN categories). These will represent the seven distinct categories of capital equipment used in the plant.

This level of aggregation is largely dictated by the level at which the data is collected and reported. In particular, the Flat Sorting Machinery category (920) includes the plant's expenditures on both the FSM881 and FSM1000 technology. Expenditures on these two types of machines cannot be distinguished. Similarly, the expenditures on Bar Code Sorters (950) includes purchases of both BCS and DBCS machinery. We have separated Bar Code Readers (950000) from the rest of the 950 category because BCR capital can represent equipment used to sort letters or flats. We will construct a separate capital stock measure for BCR and use it as an input into both letter and flat-sorting operations. The same thing will be done with the capital equipment in the final Other Capital category since this equipment can be used in either operation in the plant.

At this point the capital expenditure data have been aggregated into seven categories of capital equipment by year of acquisition. This is the book value of capital by type and year of acquisition. To construct measures of the plant's capital stock in each of the seven categories it is necessary to aggregate over the book value of equipment acquired in different years. Ideally, we would like to make two adjustments to the book value data at this point. First, we would like to deflate the nominal expenditures by the price of the equipment in the year it was acquired to convert to real expenditures. This requires investment price series that go back as far as the oldest equipment that is still in operation at each point in time. In response to an inquiry by the OCA, the USPS provided price indexes for new equipment for the period 1993-2000 (LR-J-209) but these are not sufficient to deflate expenditures made in earlier time periods. Rather than use nominal expenditures for some equipment and real expenditures for others we have chosen to not deflate any of the capital expenditure data for the plant and to simply work with the nominal expenditures.

The price index for automated equipment provided in LR-J-209 declines over time, from 1.0 in the first quarter of 1993 to .86 in the fourth quarter of 2000. This means that in aggregating over assets of different ages to construct the plant's capital stock for automated equipment categories we will underestimate the contribution of the book value of assets acquired in recent years. The price index for mechanized equipment rises over time from 1.0 to 1.17 in the period 1993:1 to 2000:4 so the error in aggregating assets of different ages will work in the opposite direction from the automated equipment.

The second adjustment we would like to make before aggregating across acquisition years is to recognize the physical depreciation of the asset as it ages. Generally this is done by collecting information on the service life of the asset and assuming a constant rate of exponential decay in value over that life. One advantage of the capital data that is collected by the USPS is that assets that are retired are removed from the plant accounts and so are not included in the plant's book value once they are retired. We will not attempt to control for physical depreciation of the assets but rather maintain the assumption that the asset maintains its original value until it is retired. If the assets do depreciate over their lifetime then we will overweight the contribution of older assets in constructing the plant capital stocks.

Overall, our estimate of each of the seven plant capital stocks at each point in time is equal to the book value of the non-retired assets in the plant. It is important to recognize that, while this is not a perfect measure of the plant's capital input, it will pick up cross-sectional differences in the size of plants, since larger plants will tend to have higher expenditures whether measured in real or nominal terms. It will also allow us to control for differences in the extent to which plants use a technology. For example, plants that are heavy users of the DBCS technology will, controlling for output level, tend to have larger capital expenditures in the BCS category than plants that have not adopted the technology. In short, differences in capital stocks across plants arising from differences in the level of expenditures on a type of capital good are large and likely to dominate any differences in capital arising from the differences in the age of the machinery in the plant. The book value of capital that we construct will reflect these large cross-plant differences in the amount of capital in place.

IV.C Sample Selection

The data analyzed in USPS-T-14 begins with quarterly observations for 321 plants. They use a number of filters to identify outlier observations, particularly plants with very high or low productivity, where productivity is defined as the ratio of TPF/manhours in each sorting operation. They lose additional observations because they use lagged values of output as explanatory variables and so for an observation to make it into the sample the plant must have positive output in the sorting operation for the current and 4 previous quarters.

We adopt a different approach to cleaning the data. In the sample there are six plants that do not report capital expenditures in many or all years and those plants are deleted from the beginning.[15] An additional 14 plants had obvious errors in either the FHP count for letters or flats or the hours data for one of the operations.[16] These plants are also deleted from the analysis and

---

[15] The plants have id numbers 17, 18, 27, 44, 196, and 197. We do not delete a plant if the plant's capital stock in one of the asset categories is zero. For example, some plants report no capital stock in the flat-sorting machinery category but also have no manhours or TPF counts in the automated flat-sorting categories (19 or 20). These are plants that do all flat sorting in manual operations. We want to include these plants in the estimation of labor demand equation for manual sorting (category 05).

[16] Generally, these plants have obvious coding mistakes in the FHP or hours data. Often this will be a single quarterly observation that is a factor of ten larger than the adjoining time period observations or similar coding errors. Many of the observations for these plants could be included in the analysis but, because of the time constraint, we chose to eliminate the plant entirely. The plants that were deleted have id numbers: 33, 35, 41, 42, 46, 54, 68, 121, 133, 142, 144, 198, 201, and 217.

we were left with a set of 301 plants to analyze. The observations are quarterly for the years 1994-2000, a total of 28 observations per plant, and this gives a maximum of 8428 observations for each sorting operation. Unlike the analysis conducted by the Postal Service, we do not use the data on TPF or TPH by MODS category and so we do not impose any selection criteria based on those variables.

To account for the phase-in or phase-out period of a technology, we drop all observations within one year (4 quarters) of the introduction or elimination of a technology in the plant. This affects labor categories 02, where the LSM technology was being phased out, and categories 17, 18, and 20, where the FSM881, FSM1000, or DBCS technologies were being phased in. These observations were deleted because we observed that during the initial introduction period the manhours used in the category were often very erratic. It took approximately a year before the hours stabilized and the technology appeared to be fully integrated in production.[17]

V. Econometric Methodology

To estimate labor demand models using a plant-level panel data set we must address a number of potential econometric issues. Because the goal of the empirical model is to accurately estimate the variablility of labor demand with respect to output, the most crucial econometric problems are ones that lead to correlation between the output variable and the error term in the labor demand equation. To frame the discussion we will use the simplified regression equation:

---

[17] In the Postal Service analysis they eliminate observations for a labor demand category if the four previous quarters do not have a positive TPF count in the operation. This has the implication that, like the criteria we use, it eliminates observations corresponding to the first year of introduction of a new technology in the plant.

(6) $$h_{it} = \beta X_{it} + \gamma\, q_{it} + \varepsilon_{it}$$

The dependent variable h is the natural logarithm of manhours in one of the sorting operations. The subscript it represents a single observation for plant i in time period t. X is a vector of explanatory variables including the capital stocks, relative wages, technology dummies, time trend, and quarterly dummies discussed in section III above and β is a corresponding vector of parameters to be estimated. The key explanatory variable of interest is q, the natural log of output, which is why it is singled out in this equation and not included in the X vector. The output variability of labor demand is the parameter γ which is the main focus of interest in estimation.[18] The random error term ε captures all factors besides the variables in X and q that lead to variation in h.

There are many ways to estimate the parameters β and γ from a sample of data on mail-sorting plants and the appropriate method depends on the behavior of the random variables $X, q,$ and ε.[19] Under the conditions that $E(\varepsilon) = 0$ and $Cov(X,\varepsilon) = Cov(q,\varepsilon) = 0$ then the ordinary least squares (OLS) estimator provides consistent estimates of the parameters. The random variable ε must have a zero mean and be uncorrelated with each of the explanatory variables in the model. The important issue for applied researchers is whether or not these conditions are likely to be satisfied in their application and, if they are not, what alternative estimators have better statistical properties than OLS. The zero mean assumption is satisfied if the regression model contains an intercept so the important conditions to verify are the zero covariances between the error term and

[18] This parameter is usually referred to as the output elasticity of labor demand in economic studies.

[19] The issues to be discussed in this section are covered in detail in virtually all econometrics textbooks. The discussion here will follow the presentation in Jeffrey Wooldridge, *Econometric Analysis of Cross Section and Panel Data,* MIT Press , 2002.

the explanatory variables. If $Cov(X_1, \varepsilon) = 0$ then we say that the explanatory variable X_1 is *exogenous* and if $Cov(X_1, \varepsilon) \neq 0$ then X_1 is *endogenous*.

The crucial econometric issue for this study is whether the output variable q is exogenous in the labor demand equations. If output is exogenous or, equivalently, $Cov(q , \varepsilon) = 0$ then OLS will provide consistent estimates and is an appropriate methodology to use to estimate the output variability parameter γ. If $Cov(q, \varepsilon) \neq 0$ then output is endogenous and using OLS to estimate the labor demand equations will result in biased estimates of the output variability. Virtually all of the remaining discussion in this section of the paper is focused on determining whether or not output is an exogenous variable in the labor demand equations and discussing alternatives to OLS estimation in situations where output is endogenous.

We begin with a discussion of the reasons why output may not be an exogenous variable. Wooldridge (2002, p. 50-51) describes three general ways that the exogeneity condition ($Cov(q , \varepsilon) = 0$) can be violated. The first is due to omitted variables in the regression equation. If there are systematic factors that determine the plant's labor demand that are not included in the set of explanatory variables X and q then these are captured by the error term ε. In addition, if some of these omitted factors are correlated with the plant's output level then $Cov(q, \varepsilon) \neq 0$, plant output will be endogenous in the labor demand equation, and OLS will not be a consistent estimator of the output elasticity γ. This is a potentially important issue in the labor demand equations and we will discuss its implications in more detail below.

The second way that output can be endogenous in the labor demand equation is if there is measurement error in the output variable. As described in the theoretical model in Section II, the conceptually correct output variables are counts of the number of sorted letters, flats, and parcels that are sent out of the plant. In practice we will use counts of FHP which, in turn, are estimated from the weight of mail received and national conversion factors which themselves have changed over the sample period. Deviations between the conceptually appropriate output variables and the FHP variables used in estimation are captured by the error term ε. Under certain circumstances this can lead to endogeneity of output and is an issue that must be addressed in the labor demand models.

The third potential source of endogeneity of the output variable arises if manhours of labor in a sorting operation and output are chosen simultaneously in the plant. This is likely to be the case if the output measure is specific to a sorting operation. For example, if the dependent variable h is manhours in MODS operation 19 (FSM881) and the output variable q is a count of pieces sorted (FHP, TPF, or TPH) *in that operation* then it is likely that h and q are simultaneously chosen. The theoretical model developed in Section II implies this would be the case. Given a flow of mail to be sorted, the plant manager simultaneously allocates the pieces across sorting operations and chooses the amount of labor to use in each operation given the capital stocks, relative wages, and other characteristics of the plant. This simultaneous choice of hours and output for each sorting operation implies that random factors (captured in ε) that lead the plant manager to vary hours will also result in variation in output in the operation. In other words, if output is defined at the level of the sorting operation, then it is likely that output is endogenous and $Cov(q, \varepsilon) \neq 0$. The empirical model estimated in USPS-T-14 defines output at the level of the sorting operation and it is likely that the exogeneity condition will be violated in this framework.

In contrast, the model that will be estimated in this paper defines output as FHP counts for the whole plant, not for specific sorting operations. At the plant level, the flow of mail received is determined by the characteristics of the mail users in the region served by the plant and not by the staffing decisions of the plant manager. Random factors that lead to variations in manhours in a sorting operation appear unlikely to have a simultaneous effect on the total volume of letters, flats, or parcels received by the plant in the quarter. This is what we meant in Section II.A when we said that the flow of mail into the plant and the sorted number of letters, flats, and parcels produced in the plant were exogenous. The implication of this for our econometric estimates of γ is that, because we use FHP counts as the output variables, this third factor that could contribute to a violation of the condition $Cov(q , \varepsilon) = 0$ will probably not be a problem in this study.

Overall, when estimating the labor demand models from section III using plant level data it will be important to address the potential endogeneity of the output variable arising from the first two factors, omitted variables and measurement error in output. It should not be necessary to be concerned about endogeneity arising from the third factor, the simultaneous choice of hours in a particular sorting operation and the total output of the plant. Because the labor demand models estimated by the Postal Service define output at the level of the sorting operation, it is possible that this third source of output endogeneity is also a problem in their framework. In the next two sections we will examine the implications of omitted variables and output measurement error for the estimation of the output variabilities.

V.A Omitted Variables

In production and factor demand models that are estimated with micro data it is difficult to control for all variables that may affect the plant's output or input choice. It has been long recognized in the applied production literature that it may not be possible to measure all the inputs used by a plant. In particular, the input of managerial expertise is often impossible to measure. If plants differ in the amount or quality of the managerial input but it is impossible to measure then this input acts like an omitted variable that is captured in the error term of the labor demand equation. Denote the omitted managerial input as m , and assume for simplicity that m is uncorrelated with all of the explanatory variables in X but is correlated with the output level q. Under these conditions the OLS estimator $\hat{\gamma}$ will be a biased estimator of the output elasticity γ.

As shown in Wooldridge (2002, p. 62), the asymptotic bias of the OLS estimator can be written as

$$plim\ \hat{\gamma} - \gamma = \beta_m\ (Cov(q,m) / Var(q)),$$

where β_m is the true effect of the omitted managerial input on labor demand, $Cov(q,m)$ is the covariance between the included output variable and the omitted managerial variable, and $Var(q)$ is the variance of output. This bias, the right side of the equation, can be either positive or negative. The OLS estimator can overestimate or underestimate the true variability γ depending on the signs of β_m and $Cov(q,m)$. A common interpretation of these terms would have $\beta_m <0$, because the omitted managerial input would be a substitute for other kinds of labor and thus reduce other labor demands, and $Cov(q,m) >0$ because plants with more of the managerial input would tend to be larger.[20] In this case the OLS estimator would underestimate the true output variability. In effect, the labor-saving contribution of this managerial input would be incorrectly attributed to output, so we would underestimate the effect of output changes on labor demand.[21]

In the labor demand equations for the postal plants the sign of the bias arising from potential omitted variables is difficult to determine a priori. To do so we must be specific about the likely omitted variables, their partial effect on the plant's labor demand holding output, capital,

[20] This would be a realistic argument if the plant produced and sold its output in a competitive market. Plants with more or better managerial input would tend to be more successful and grow larger than their less gifted rivals in a competitive environment.

[21] It is important to recognize that, while an estimator may be a biased estimator of a particular parameter, the magnitude of the bias may be small. For example, if the omitted managerial input has little effect on the plant's demand for labor , so that β_m is close to zero, or the omitted managerial input is not strongly correlated with output, so that $Cov(q,m)$ is close to zero, the bias term will be close to zero

wages and other observed characteristics fixed, and their correlation with output. It may be that the explanation in the previous paragraph, which is based on an omitted managerial input, is correct, in which case OLS estimates of the output variability would be too small. In contrast, an alternative explanation may be that the omitted variable m is not a managerial input but rather a source of inefficiency in the plant such as a measure of organizational slack, down time, or coordination problems that exist. If the presence of this source of inefficiency means that more labor input is required to sort a given volume of mail then $\beta_m > 0$, and if it is more common in large plants then $Cov(q,m) > 0$. In this case the bias of the OLS estimator of γ is positive, meaning that OLS will overestimate the true output variability. In effect, the labor-using or labor-wasting effects of this organizational inefficiency are incorrectly attributed to output and the effect of output changes on labor use is thus overestimated.

A final example reflects that the omitted variable bias could arise from differences in the skill of the workers across plants. If skill can act as a substitute for manhours, that is, more skilled workers can produce the same output as a less-skilled group but with fewer manhours, then we would like to control for differences in worker skills across plants when estimating the demand for manhours. If this is unobservable, differences in skills would act like an omitted variable in the labor demand equation. We would expect $\beta_m < 0$, since manhours and skills are substitutes. If larger plants had more skilled workers then $Cov(q,m) > 0$ and the OLS estimator would underestimate γ, while if skilled workers were more common in small plants $Cov(q,m) < 0$ and OLS would overestimate γ. The main point is that there are many reasons to expect that omitted variables bias will be present in OLS estimates of the output elasticity, but it is not possible to sign the direction of bias without a clear understanding of the omitted variable and its relationship with output.

The simple way to correct for omitted variables bias is to measure the relevant variable and include it in the X matrix of explanatory variables. Of course, in many situations the desired variable cannot be measured given the existing data and this solution is not feasible. An alternative solution to the omitted variables problem is feasible in our case because panel data, repeated time-series observations on the set of mail sorting plants, is available. In this case, we can control for all omitted factors that vary across plants but are constant over time for a each plant. Denote the variable m_i to represent all factors that are specific to plant i but that do not vary over time for the plant. In the generic labor demand equation (6), the m_i are included as part of the error term but now we specify them separately in the demand equation. This gives the demand equation:

$$(7) \qquad h_{it} = \beta X_{it} + \gamma q_{it} + m_i + \varepsilon_{it}$$

We will refer to m_i as a plant-specific effect. It is important to recognize that the plant-specific effect is not an observable or measureable variable. Rather, it represents all time-invariant factors that lead plant labor demands to differ, after controlling for X and q. For this reason it is often referred to unobserved heterogeneity, since it leads to heterogeneity in the labor demand equations across plants. Estimating equation (7) without controlling for the presence of the plant-specific effect m_i , which is what OLS estimation of equation (6) does, will lead to omitted variables bias in γ if the plant-effect is correlated with the output variable. Since output tends to vary substantially across plants, it is very likely that the two will be correlated.

Even though m_i is not observable we can still control for its presence in equation (7) when we have panel data. The presence of multiple time observations for each plant i allows several ways to control for the presence of m_i when estimating the labor demand curve. One alternative is to difference equation (7) across different time periods. Rewrite equation (7) for period *t-1:*

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

All the time-varying variables, *h, X, q*, and ε have different values in period *t-1*, but the plant-specific effect does not. Taking the difference between period *t* and period *t-1,* (subtracting equation (7') from equation (7)), gives the labor demand equation written in terms of changes over time:

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

All variables are now expressed as differences between their value in period *t* and period *t-1*.[22] Because m_i is constant over time the differencing eliminates it from the equation. Since the omitted variable m_i is the reason the output variable is endogenous, the time differencing of the equation removes the source of endogeneity. The parameters of equation (8) can be consistently estimated without worrying about omitted variables bias arising from the plant-specific effect. When we estimate the parameters β and γ using equation (8) we will refer to estimating the model in first differences or say that we are using the First Difference (FD) estimator.

A second way to eliminate the bias caused by the presence of the plant-specific effect is to treat m_i as a plant-specific parameter that is estimated. This can be done by including a set of

[22] For variables like *h* and *q* that are expressed as natural logarithms the differences are the period-to-period growth rates in the variable.

dummy variables, one for each of the plants in the data set. This is why the m_i is often referred to as a plant-specific intercept. An equivalent method is to express each plant's data (*h, X,* and *q*) as deviations from the plant-specific mean of the variable. In this case the estimating equation becomes:

$$(9) \qquad h_{it} - \overline{h}_i = \beta\left(X_{it} - \overline{X}_i\right) + \gamma\left(q_{it} - \overline{q}_i\right) + \left(\varepsilon_{it} - \overline{\varepsilon}_i\right)$$

where the variables with bars above them are the plant-specific means. Since the m_i does not change over time for a plant it is equal to the plant-specific mean value and, once again, the differencing of the data removes it from the equation. Once the estimating equation has been expressed in this form the m_i is eliminated and the source of output endogeneity is removed. When we estimate the parameters β and γ using equation (9) we will refer to this as the Fixed Effects (FE) estimator. Both the FD and FE estimator correct for the omitted variable problem in the case where the omitted variable differs across plants but is constant over time for each plant.

When using micro data on individual mail-processing plants it is likely that some factors that affect the labor use in a plant are omitted from the estimating equations. A partial list includes: the skill or experience level of the work force, the motivation and effort of the employees, the level of absenteeism or job turnover, the degree to which workers specialize in specific tasks, the number or quality of supervisory personnel, the physical size of the plant, the layout of the capital equipment in the plant that affects the flow of mail through the sorting process, the amount of presorting that is done to the mail arriving at the plant, or the depth of sorting that is done within the plant. Factors like this are captured in the error term in the regression equations and, if they are not systematically related to the output level of the plant, do not lead to bias in the OLS estimates of the output elasticities. If they do vary systematically across plants of different sizes then their presence will bias OLS estimates of the output elasticities. The inclusion of a plant-specific effect in the labor demand equation and the resulting use of either FE or FD estimators will control for across-plant differences in these characteristics and remove the bias from the estimates of the output elasticity.

Following estimation it is possible to assess the importance of the omitted variables problem and the ability of the model with plant-specific effects to control for it. After estimating the plant-specific intercepts we can test the hypothesis that they are jointly equal. That is, we can test the null hypothesis $m_1 = m_2 = ... = m_n$ for the n plants in the sample. Rejection of this hypothesis indicates that there are systematic differences across plants in the number of manhours used in a sorting operation after controlling for the observable plant characteristics. If those differences are correlated with the size of the plant, then failing to correct for them would lead to biased estimates of the output variability.

V.B Measurement Error in Output

As indicated in the beginning of section V, a second potential source of endogeneity in the output variable is measurement error. In our case this arises from the fact that the observed FHP counts of letters and flats are not exact measures of the number of unique letters and flats sorted in the plant. To see how measurement error can arise in the construction of the FHP variable we need to be specific about how it is constructed. Let Q_{it} be the measured number of FHP letters of a particular category that we observe for plant i in time period t. That is constructed as the product of the measured weight in pounds of letters received in the plant (W_{it}) and a constant national conversion factor δ that is measured as letters per pound for that category of mail.

(10) $$Q_{it} = W_{it}\ \delta$$

The measured weight W_{it} is equal to the true number of letters received $\left(Q_{it}^{*}\right)$ multiplied by the true average weight (pounds per letter) of these letters $\left(AW_{it}^{*}\right)$. The variable that we would like to measure and include in the labor demand equations is Q_{it}^{*} the true number of letters received in the plant, but this is not observable. Substituting $W_{it} = Q_{it}^{*} AW_{it}^{*}$ into equation (10) gives an expression for the observed FHP count Q_{it} in terms of the true unobserved output Q_{it}^{*}.

(11) $$Q_{it} = Q_{it}^{*} AW_{it}^{*} \delta .$$

Rewriting this in logs gives:

(12) $$q_{it} = q_{it}^{*} + aw_{it}^{*} + ln\delta$$

where the lower case letters are used to denote the logarithm of the corresponding variable. This equation shows that the output variable used in the labor demand equations q_{it} can be written as the sum of the true unobserved output measure q_{it}^{*} and an error, aw_{it}^{*}, that reflects the average weight of letters in the category in the plant.[23] If the variables q_{it}^{*} and aw_{it}^{*} are uncorrelated then equation (12) shows that the log of observed FHP letters satisfies the conditions of the classical errors-in-variables model: the observed variable can be written as the sum of the true unobserved variable and an uncorrelated measurement error.[24] In the case at hand it is reasonable to treat q^* and aw^* as uncorrelated variables. There is no strong reason to expect that bundles of mail composed of many pieces contain systematically heavier or lighter pieces. It is reasonable to assume that the weight of a bundle varies because there is variation in both the total number of pieces in the bundle and the average weight of the pieces in the bundle but the two sources of randomness are unrelated to each other.

Equation (12), combined with the assumption that q^* and aw^* are uncorrelated, implies that the observed output variable used in the regressions q will be endogenous. To see this note that the general regression equation (6) should be expressed with the true output variable q^* as the explanatory variable:

$$(13) \qquad h_{it} = \beta X_{it} + \gamma q_{it}^{*} + \varepsilon_{it}$$

This is the equation we would like to estimate but q^* is not observable. Solving equation (12) for q^* and substituting it into equation (13) gives the labor demand equation in terms of the observed output variable q:

---

[23] The constant $ln\delta$ will simply be absorbed into the intercept of the regression equation and is of no concern.

[24] See Wooldridge (2002), p.73-76 for discussion.

(14) $$h_{it} = \beta X_{it} + \gamma\, q_{it} - \gamma\, ln\, \delta + (\varepsilon_{it} - \gamma a w_{it}^{*})$$

The error term in this equation is in parentheses and consists of two components: the purely random shocks to hours represented by ε_{it} and the error in the measured output variable represented by $\gamma\ aw_{it}^{*}$. This latter term is correlated with the observed output variable, as shown in equation (12), and makes the observed output variable endogenous in (14). .

Because of the endogeneity of *q,* OLS estimates of equation (14) will produce biased estimates of the output elasticity *γ.* In the case where *q* is the only explanatory variable in the labor demand equation, the asymptotic bias of the OLS estimator of *γ* can be represented with the equation:

(15) $$plim\ (\hat{\gamma}) = \gamma\ [\ Var(q^*) / (\ Var(q^*) + Var(aw^*)].$$

This has a very simple interpretation. The expression in square brackets is the variance of the true unobserved output variable *q** relative to the variance of the observed output variable *q*. This ratio will always be less than one so the OLS estimator of *γ* will be biased toward zero. In our case, since the true *γ* is positive, OLS will underestimate it.[25]

[25] This intuition will also hold if there are additional explanatory variables *X* in the labor demand equation. As shown by Wooldridge (2002, p. 75, eq. 4.47) the plim $(\hat{\gamma}) = \gamma\ [\ Var(e) / (\ Var(e) + Var(aw^*)]$ where *e* is the linear projection error from a regression of *q** on the explanatory variable matrix *X.* This variance ratio will still be less than one indicating that OLS will underestimate *γ* if the true value is positive.

As with omitted variables bias, it is reasonable to ask how large the measurement error bias may be. From equation (15) we can see that if the variance in the error aw^* is small relative to the variance in the true output level q^* then the term in brackets will be close to one and the bias in the output elasticity will be small. Unfortunately, it is not possible to check if this is true by examining the data since both q^* and aw^* are unobserved. However, we do have some indirect evidence on how aw^* has changed over time. As shown above, aw^* is the true average weight (pounds/piece) of mail in a weighed bundle. This is simply the (inverse) of the true conversion factor that would allow us to move between the observed total weight and the true number of pieces of mail in the bundle. It will differ for every bundle of mail. The USPS estimates conversion factors for several categories of mail for the country as a whole and, as discussed in section III.E, these conversion factors have changed over time, sometimes by a substantial amount. Based on the calculations in section III.E, we estimate that the change in conversion factors at the beginning of 1999 contributed to approximately an 18 percent decline in measured FHP letters and an 11 percent decline in FHP flats. This suggests that, over time, the true average weight of letters and flats had risen significantly. While this does not provide any evidence on the cross-plant variation in aw^*, it does suggest that at some points in time there may have been significant errors in constructing FHP counts based on measured weights and the use of national conversion factors.

V.C Instrumental Variables Estimator

In the last two sections we have discussed two situations, omitted variables and measurement error, that will lead the output variable q to be endogenous in the labor demand equation (6). Exploiting the panel nature of the data to incorporate plant-specific intercepts and using FD or FE estimators can eliminate the omitted variables bias. Unfortunately, the use of FD or FE estimators in the presence of measurement error can actually increase the bias due to measurement error.[26] The intuition behind this result is that both FD and FE estimators rely on transformations of the data that remove some of the cross-section variation among plants (because that is the primary source of the omitted variables problems) and thus rely more heavily on the time series variation in the data to identify parameters. If the cross-sectional differences in plant output are large relative to the time-series differences for the individual plants, then these transformations of the data will tend to reduce the systematic variation in the data relative to the measurement error. If there is measurement error in the output variable, then FD or FE estimates of the output variability γ are likely to be more substantially affected (biased toward zero) by the measurement error than OLS estimates.[27] As a result it is important to deal with both the omitted variables and measurement error problems in a unified way in order to consistently estimate the output variability.

In docket R2000-1, there was testimony presented by the Postal Service (USPS -T-15 and USPS-RT-7) and by United Parcel Service (UPS-T-1) and discussion in the Commission's final

---

[26] Zvi Griliches, "Economic Data Issues," in *Handbook of Econometrics*, Z. Griliches and M. Intriligator (eds), Elsevier Science Publishers, 1986 and Zvi Griliches and J.A. Hausman , "Errors in Variables in Panel Data," *Journal of Econometrics*, Vol 31, 1986, p. 93-118, discuss this issue in detail. They also show how instrumental variables methods can be used to solve the problems that arise in panel data.

[27] It is common in the applied production literature to observe that differencing the data, as FD or FE estimators do, results in substantially smaller coefficients than the OLS estimator. Examples of this pattern are discussed in Griliches, Zvi and Jacques Mairesse, "Productivity and R&D at the Firm Level," in Z. Griliches (ed.)*, R&D, Patents, and Productivity* (University of Chicago Press), 1984, Mairesse, Jacques "Times Series and Cross-Sectional Estimates on Panel Data: Why Are They Different and Why Should They Be Equal?", in J. Hartog, G. Ridder, and J. Theeuwes (eds.), *Panel Data and Labor market Studies* (Elsevier Publishers), 1990, Mairesse, Jacques and Brigitte Dormont, "Labor and Investment Demand at the Firm Level: A Comparison of French, German, and U.S. Manufacturing, 1970-1979, *European Economic Review*, Vol 28, 1985, Tybout, James R. and M. Dan Westbrook, "Estimating Returns to Scale with Large Imperfect Panels: An Application to Chilean Manufacturing Industries," *World Bank Economic Review*, Vol 7, 1993, and Roberts, Mark J. and Emmanuel Skoufias, "The Long-Run Demand for Skilled and Unskilled Labor in Colombian Manufacturing Plants," *The Review of Economics and Statistics*, Vol 79, May 1997.

decision concerning the role of measurement error, the biases that it could induce, and whether or not it was likely to be a problem in this application. There was no discussion, however, of how the problem could be corrected. In the remainder of this section we will (1) propose a methodology that can correct for the inconsistency caused by measurement error, (2) discuss how to apply it to the labor demand equations, and (3) show that it also provides a basis for testing whether the output variable is endogenous in the estimating equations.

As discussed in the beginning of section V, the basic econometric problem that arises with either omitted variables or measurement error in output is that $Cov(q, \varepsilon) \neq 0$ so that the output variability cannot be estimated consistently by applying OLS to equation (6). One solution to this problem is provided by the method of instrumental variables (IV).[28] If there exists an observable variable z that is correlated with the output variable q but uncorrelated with the regression error ε then it will be possible to use the variable z to correct for the fact that q is endogenous. More specifically, z must be a variable that is not included among the explanatory variables in equation (6), that is, it is not a determinant of the labor demand equation being estimated. It must satisfy two additional properties. First, it must be uncorrelated with the error term in the labor demand equation, $Cov(z, \varepsilon) = 0$. Second, it must be correlated with the endogenous variable q after controlling for the other explanatory variables X in the model. Specifically, in the regression $q = \alpha X + \theta z + w$, the coefficient θ must be nonzero. This indicates that z has a role to play in explaining variation in q even after controlling for all the other variables in the model. A variable z that satisfies these two requirements will be referred to as an instrumental variable or, more simply, an instrument for q.

[28] This discussion follows Wooldridge (2002), chapter 5, particularly sections 5.1 and 5.3.

The instrumental variables (IV) estimator is formally defined in most econometrics textbooks, for example Wooldridge (2002, p. 86), and we will not present the formula here. It is possible to write the IV estimator as resulting from a two-stage estimation process. In the first stage, the endogenous variable is regressed on the exogenous variables X and the instrument z using OLS, and the fitted value of the endogenous variable is constructed. In our application, we construct

(16) $$\hat{q}_{it} = \hat{\alpha} X_{it} + \hat{\theta} z_{it}$$

where $\hat{\alpha}$ and $\hat{\theta}$ are the OLS estimates. In the second stage, q_{it} is replaced with $\hat{q}_{it}$ in the labor demand equation:

(17) $$h_{it} = \beta X_{it} + \gamma\ \hat{q}_{it} + \varepsilon_{it}$$

and this equation is estimated with OLS. The resulting estimates of β and γ are the IV estimates.

The IV estimator is a consistent estimator of γ when $Cov(q, \varepsilon) \neq 0$ although it will be biased in small samples. When using the IV estimator there is an important practical consideration that must be addressed. In the first stage regression, equation (16), it is important that the instrument z have some explanatory power. If the partial correlation (i.e. after controlling for X) between the endogenous variable q and the instrument z is low then the small sample bias of the IV estimator can be substantial and the asymptotic variance of the IV estimator can be large (Wooldridge 2002, p.101-104). The latter can result in large standard errors for the IV estimates, particularly for the coefficient on the endogenous variable, and make it difficult to reject hypotheses that are tested.

Besides providing consistent parameter estimates in the case where output is endogenous, the IV estimator also provides an opportunity to test whether $Cov(q, \varepsilon) = 0$. If we cannot reject this hypothesis then it indicates that the omitted variables or measurement error problems that create the endogeneity are not a concern and helps justify the use of simpler OLS estimators. The formal method for testing for endogeneity is due to Hausman (1978) and involves a comparison of the OLS and IV estimator.[29] If output is truly exogenous in the labor demand model, the OLS and IV estimators will differ only due to sampling variation and Hausman presents a formal test statistic that the OLS and IV estimators are identical. The test is simple to implement in our framework. First, estimate the same first-stage regression as used in constructing the IV estimator. That is, regress q_{it} on X_{it} and the instrument z_{it} using the OLS estimator. Construct the residuals $e_{it} = q_{it} - \hat{\alpha}X_{it} - \hat{\theta}z_{it}$. Second, include the residuals as an additional explanatory variable in the original labor demand equation and estimate it with OLS. That is, estimate the equation

$$h_{it} = \beta X_{it} + \gamma q_{it} + \lambda e_{it} + \varepsilon_{it} \quad (18)$$

using OLS. Under the null hypothesis that q is an exogenous variable, the coefficient $\lambda=0$, and this can be tested using a t-test on the estimated $\hat{\lambda}$ from equation (18). If we reject the hypothesis that $\lambda=0$ then we reject that q is exogenous and must conclude that the conditions necessary to justify the use of OLS estimators of the labor demand model are not satisfied. IV methods are

[29] Hausman, J.A. "Specification Tests in Econometrics," *Econometrica,* Vol. 46, p. 1251-1271. For further discussion see Wooldridge (2002), p. 118-122.

needed to correct for the endogeneity of output.

The IV method as outlined in this section is very general and can be applied in many situations where a variable is potentially endogenous. The important practical consideration is to construct a variable, or set of variables, that meet the conditions required for an instrument. To see how this can be used to correct for measurement error in output we focus on equation (14). The endogeneity arises because q is correlated with the measurement error aw^*. In the case of a labor demand equation for a letter sorting operation (MODS categories 01, 02, 06, 17 , and 18) q is the measured number of FHP *letters* processed in the plant and the measurement error arises from variation in the average weight of letters across plants and time periods. The instrument must be correlated with the true (unobserved) number of letters handled in the plant but be uncorrelated with the average weight of letters in a plant. A variable that will satisfy these requirements is the number of FHP *flats* processed in the same plant and time period. The number of flats and letters will be correlated across plant-time observations because of cyclical patterns in mail volume, systematic differences in the size of the plants, and characteristics of the population in the area served by each plant. When estimating demand equations for letter-sorting operations we will use the count of FHP flats to instrument the error-ridden output variable.[30] Similarly, when estimating demand equations for flat-sorting operations the output variable is the number of FHP flats and we

[30] It does not matter that the observed count of FHP flats is itself an error-riddden measure of the true (unobserved) number of flats processed in the plant. All that matters is that the observed variable for FHP flats be correlated with the true unobserved number of letters sorted in the plant and uncorrelated with the measurement error in letters. This will be satisfied if the measurement errors in FHP flats and FHP letters are independent. This is very likely to be the case since the sources of measurement error in the two variables arise from entirely different sources. One arises from the fact that the average weight of letters for an observation varies from the national conversion factors for letters while the other arises because the average weight of flats varies from the conversion factor used for flats.

will use the count of FHP letters as the instrumental variable.[31]

Using the IV estimator we can address both the omitted variables and measurement error problems in a single, consistent way. As described above, the omitted variables problem can be addressed with the use of FD estimator of equation (8) or the FE estimator of equation (9). In either case the output variable is now a deviation, either a deviation over time or a deviation from the plant-specific average output level, but this variable is still subject to measurement error problems. It is necessary to find an instrument for the output *deviation* that appears as the explanatory variable in (8) and (9). In each case we will use the corresponding deviation of the other FHP output variable in the plant. For example, when using (8) to estimate labor demands for letter sorting operations, we will use the change in FHP flats as the IV for the change in FHP letters. In (9) we will use the FHP flats expressed as a deviation from the plant's mean value as an IV for the deviation in FHP letters. Similarly, we will use the deviation in FHP letters as an IV for the deviation in FHP flats when estimating demands in flat-sorting operations.[32]

---

[31] It is possible to use more than one instrument for each endogenous variable. One additional candidate for an instrumental variable is a measure of the capital stock used in the *other* sorting operations. For example, the capital stock of flat-sorting machinery (KFSM) could be used as an IV for FHP letters. (Note that KFSM would not be a valid instrument for FHP flats since it is already included as an explanatory variable in the labor demand equations for flat-sorting operations). As we will document below, the output variables are good instruments and we found that adding further measures of capital added very little explanatory power to the instrument set and so this avenue was not pursued further. Other candidates for instrumental variables are lagged values of the endogenous variable. That is, the plant's count of FHP letters in previous periods could be used as IV's for the current period count of FHP letters. The use of lags as instruments in FD or FE models generally produces very unsatisfactory results because of the low correlation between changes in the endogenous variable and the level of the instrument. This can result in small sample biases that are substantial and large standard errors for the IV estimates. Blundell, Richard and Stephen Bond, "GMM Estimation With Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, Vol 19, 2000 discuss this problem and suggest a methodology that can allow the researcher to exploit lagged variables as instruments. In our application the use of the other output variable in the plant as an instrument works quite well in the models we estimate and we did not utilize lagged endogenous variables as instruments. In the future, if we are able to develop good measures of the FHP count of parcels and priority mail in the plant, they could be incorporated as additional instruments for both letters and flats.

[32] This general approach of differencing the data to remove plant-specific effects and then applying an IV estimator to correct for an endogenous variable is discussed in detail in Wooldridge (2001), p.305-315. The IV estimator can also be used to correct for the endogeneity resulting from an omitted plant effect, but it will generally require a different type of instrument. For example, focusing on the model with an omitted plant effect (equation 7) where m_i is

V.D Comparing Alternative Estimators of the Demand for Manual Labor in Flat Sorting

Estimation of the relationship between output and employment is complicated by the different mix of technologies across observations (time and plants) and the likely importance of capital as a determinant of the plant's use of manhours. To isolate and illustrate the importance of the econometric issues that have been identified in this case, and to demonstrate the role that IV can play in the estimation of labor demand functions, we focus on a simplified case that is present in the data. In the data set there are a total of 45 plants that do all of their flat sorting in manual operations. These plants report no capital equipment for flat sorting, no hours in any of the mechanized MODS categories (18 or 19), and no TPF in either operation 18 or 19. All flat sorting is done with manhours in a single MODS category (05). For these plants, the labor demand function derived in section III.H simplifies so that it only depends on the output, trend, quarterly dummy variables, and the dummy variable denoting the change in conversion factors. Flat sorting machinery, the presence of other technologies for sorting flats, and the relative wage for automated vs manual manhours do not matter to the plant's use of manual labor. In order to illustrate the importance of the econometric issues, and the effect that the corrections have on the estimated output elasticity, we will use the data for this, admittedly, specialized group of plants and estimate simple log-linear models of manhours in manual flats sorting. The explanatory variables will be the log of FHP flats, a dummy variable to control for the change in output definition starting in 1999 *(DC)*, a set of three quarterly dummy variables *(DQ2, DQ3, and DQ4)*, and a time trend

---

included in the regression error, we need an instrument that is correlated with the endogenous output variable q but *not* correlated with the omitted plant effect m. If the labor demand equation was for a letter-sorting operation, so that q was the count of FHP letters, then the FHP flat variable would not be a good instrument because it would likely be correlated with m for the same reason as FHP letters. Measures of the other outputs in the plant would not be good instruments if the endogeneity of q arises because of an omitted plant effect. We will instead deal with omitted variables problems by using FD or FE estimators.

(TREND).

The output variability estimates for five different estimators are reported in Table 3. The first row reports that the ordinary least squares estimate (standard error) is .926 (.014). It is statistically different than 1.0 indicating that there are slight increasing returns to labor in this sorting operation. The OLS estimator could potentially be subject to an omitted variable bias, which could be either positive or negative, and a measurement error bias that is negative. The main alternative estimator discussed in the R2001-1 rate case is the FE estimator. In this operation, the FE estimate is .895 (.022). The hypothesis that there is no plant-specific effect in the demand equation can be tested by testing if the plant-specific intercepts are equal: $m_1 = m_2 = m_3 = \ldots = m_{45}$ using an F-test. The test statistic is 71.7 and indicates that we reject that the plant-specific effect is not important at the .01 significance level.

While this FE estimator can correct for the omitted variables bias, it is subject to a downward bias if there is measurement error in the output variable. The alternative estimation method that also corrects for the omitted variable bias by removing the plant-specific intercept from the error is the First Difference (FD) estimator. The FD estimate is .931 (.023) and is also significantly different than one. This FD estimator may also be subject to a downward bias due to measurement error in output.

Before reporting the IV estimates, it is useful to assess the explanatory power of the instrumental variable in the first-stage reduced-form regression. When using the FE estimator, the reduced form regresses the endogenous variable *lnF* on *DC, DQ2, DQ3, DQ4*, *TREND,* a set of plant dummies, and the instrumental variable *lnL*. The coefficient (standard error) on *lnL* is .764 (.016) giving a t-statistic of 46.70. The count of FHP letters is highly significant in the reduced-form regression, even after controlling for the plant effects. The overall fit of the reduced-form regression is $R^2 = .93$. When using the FD estimator the reduced form regresses the growth in *lnF* on quarterly dummies, a dummy for the observation for the time period 98:4-99:1, and the growth in *lnL*. The coefficient on the growth in *lnL* is .869 (.011) giving a t-statistic of 76.81. The overall R^2 for the reduced-form regression is .90. In both cases the count of FHP letters is highly correlated with the endogenous variable, after controlling for the other variables in the demand equation, and thus satisfies one of the requirements for a good instrument. The fit of the equation for the growth in *lnL* is particularly important because it is often difficult to find good instruments for the growth rate of an endogenous variable.

The IV estimates of the FE and FD models are reported in the last two rows of Table 3. The FE-IV estimate is .981 (.038) and the FD-IV estimate is .966 (.026). Neither estimate of the output elasticity is significantly different than one, indicating that we cannot reject the hypothesis that there are constant returns to labor in this operation. Notice that each of the IV estimates are closer to one than the corresponding non-IV estimate. This would be expected if the IV estimator removes the downward bias resulting from measurement error in FHP flats. Also, the standard errors of the IV estimates are slightly larger than the non-IV standard errors and this reduction in precision is also a common feature of IV estimators.

We can also test the hypothesis that $Cov(q, \varepsilon) = 0$ using the Hausman test. As described above in equation (18), the test is implemented by including the residuals from the reduced-form regression in the structural labor demand equation and examining the t-statistic on the estimated coefficient λ. In the FE specification, the t-statistic is -4.56 and in the FD specification the t-statistic is -2.73. In both cases we reject the hypothesis that $\lambda = 0$ and thus reject the hypothesis that the output variable is exogenous. This indicates that one of the crucial assumptions underlying OLS, FE, and FD estimators is not satisfied and that the IV estimator is needed to provide consistent estimates of the output elasticity.

Overall, in this special case it is important to control for plant-specific effects in the data and measurement error in output. The FE-IV and FD-IV estimators account for both of these potential problems. Specification tests indicate that both corrections are important. Once both problems are addressed we cannot reject the hypothesis that the output variability of labor demand in the manual flat sorting operation (for this specialized group of plants) is equal to one. Estimators that do not correct for both problems indicate there are increasing returns ($\gamma < 1$) to labor in this sorting operation.

In the next section we estimate labor demand models with a much larger set of explanatory variables. As we move to incorporate plants that use a larger set of technologies, it will be important to correct for the presence of other technologies, the relative wage rates faced by the plant, and the types and amounts of capital equipment. However, while these additions to the model will be important, they will not correct for the basic problems of unobserved heterogeneity and measurement error in output that we have seen in this section. We will continue to observe a similar pattern in variability estimates among OLS, FE, and FD estimators and continue to rely on instrumental variables methods to control for measurement error problems.

VI. Parameter Estimates for Labor Demand Equations

We estimate the eight labor demand equations developed in section III.H. Since primary interest focuses on the output variabilities we will focus on those coefficients first and report estimates using the set of econometric methods discussed in section V. After this we discuss the full set of coefficient estimates for one of the IV models. While the other coefficients are not of primary interest they are useful to examine to determine if they are sensible and consistent with institutional features of the mail sorting process.

VI.A Output Variabilities in Flat-Sorting Operations

The output variabilities are reported in Table 4. The top part of the table reports estimates for flat sorting and the bottom of the table reports estimates for letter sorting. Beginning with flat sorting, the OLS estimates indicate decreasing returns to labor in the manual operations (1.359)and virtually constant returns to labor in the overall FSM category (.980). When disaggregated into specific types of machinery the latter is a combination of decreasing returns in FSM881 operations (1.134), and increasing returns in FSM1000 operations (.551).[33]

In order to correct for unobserved heterogeneity that can bias the OLS estimates, the second two columns of the table report FE and FD estimates. In every case, the FE and FD elasticities are lower than the OLS estimates, often by a substantial margin. This pattern is consistent with a downward bias due to measurement error that is exaggerated by the correction for unobserved heterogeneity.

The final two columns of Table 4 use the count of FHP letters, or its first difference, as an

[33] The relatively small variability for the FSM1000 operation appears to reflect, in part, the fact that this technology was being phased into the sorting plants during the sample period. A common pattern that is observed for manhours in this category is that when a plant begins using this technology the manhours will increase over about 4 quarters as the technology is brought on line. During this time there is very little correlation between the quarterly flow of FHP flats in the plant and the number of manhours in FSM1000 operations. After about a year of use, the number of manhours in this operation for a plant will tend to stabilize and variations over time will then reflect the cyclical flow of plant output. When estimating the labor demand for MODS category 20, I have deleted all observations that correspond to the first year of operation of the FSM1000 equipment in the plant. Ideally it would be desirable to have a longer time period corresponding to the full-time operation of the equipment in the plant to estimate the labor-output relationship. With the introduction of the AFSM equipment in 2001, another change in technology will be introduced into future observations and it may be necessary to collect several years of data on the joint operation of these two technologies before the labor demand implications can be estimated precisely.

instrumental variable to control for the measurement error in the output of FHP flats.[34] The IV estimators perform as expected: in each of the eight cases, when compared with the non-IV estimate, the output variability rises, often by a substantial margin. We cannot reject the hypothesis that the output variability for manual operations is equal to 1.0 using either estimator. For the two mechanized operations the output variabilities differ between the estimators. For FSM881 operations, the output variabilities for the two estimators are .803 and .948. The first point estimate indicates increasing returns to labor in this operation while we cannot reject that the second estimate is equal to 1.0. The FSM1000 operations show a greater degree of increasing returns, although neither coefficient is estimated very precisely. With the FE/IV estimator the point estimate is .739 but the standard error is so large that we cannot reject the coefficient is equal to 1.0. While the FD/IV estimator is quite small, again, it is not very precisely estimated. Overall, it is difficult to precisely estimate the coefficient for the FSM1000 operation and this is very likely due to the fact that it was introduced in the middle of the sample period and more quarters of experience and data will likely increase the precision of the estimates. If the two mechanized operations are aggregated together into a single FSM-all category, the estimates from the two estimators are more similar. The FE/IV estimate is .963 and the FD/IV estimate is .916 and neither estimate is statistically different than one. Overall, the results indicate evidence of constant returns to labor in manual flat sorting and evidence of increasing returns in the mechanized operations when the mechanized operations are disaggregated by the type of machinery used.

---

[34] In the reduced-form regression for flats, the FHP letters variable has a t-statistic of 14.96. In the reduced form for the growth in flats, the growth in FHP letters has a t-statistic of 19.24. This indicates that the IV used in these regressions is significantly correlated with the endogenous variable, after controlling for other exogenous variables, and satisfies one of the requirements for a good instrument.

Table 4 reports 5 estimates for each of the four output variabilities and it is useful to conduct specification tests to distinguish among them. The first column of Table 5 reports the F-statistic for the hypothesis that all plants have the same intercept (i.e. that there is no unobserved heterogeneity). The test statistics are between 26.02 and 53.97 for the four flat-sorting operations, which indicate that the hypothesis is rejected at the .01 significance level. The last two columns report the test statistic for the hypothesis that output is exogenous in the FD or FE estimating equations. For manual sorting (05), total FSM labor (11) and the mechanized FSM881 category (19) the exogeneity of output is rejected, indicating that the FD and FE estimators are not consistent estimators of the output elasticity. With the FSM1000 operations, the exogeneity assumption is not rejected and this occurs, at least for the FE estimator, because of the large standard error on the FE/IV estimator. Overall, the specification tests provide evidence that we should be concerned about the endogeneity of output in the labor demand equations. This could arise either as a result of omitted plant effects or from output measurement errors. The specification tests indicate that either FD/IV or FE/IV estimators are to be preferred over the OLS, FE, or FD estimators that fail to correct for one or both of these problems.

In the R2000-1 rate case, there was substantial discussion about whether the variability estimates provided by the USPS analysis, which were all less than one, were more appropriate than the full volume variability assumption (γ=1) relied upon by the PRC in earlier rate cases. To contrast these alternatives we construct the probability of observing our IV point estimates under two different scenarios: the first is that the USPS estimate is the correct value of the output variability and the second is that the output variability equals one. In the first scenario we assume that the USPS estimate is the true output variability, and construct the probability of observing a point estimate at least as large as our IV estimate from Table 4. This is the p-value for the test of the null hypothesis that the true variability (γ) is equal to the USPS estimate (γ_{USPS}) against the alternative hypothesis that $\gamma > \gamma_{USPS}$. In the second scenario, we assume that the true variability equals one and calculate the probability of observing a point estimate smaller than our IV estimate from Table 4. The is the p-value for the test of the null hypothesis $\gamma=1$ against the alternative $\gamma<1$. If the p-value under the first scenario is larger than under the second, this implies that there is a larger probability of observing our point estimate if the USPS estimate, rather than one, is the true measure of the output variability. If the p-value for the second scenario is larger, there is a larger probability of observing our point estimate if the true output variability equals one, rather than the USPS estimate.

The p-values are reported in Table 6. The first column of the table reports the point estimate for each sorting operation from USPS-T-14. The second and third columns report the p-values for the null hypothesis that the USPS value is the true value using the FE/IV and FD/IV estimator, respectively. The last two columns report the p-values for the null hypothesis that the true value is one. For the manual and FSM881 operations the p-values are higher under the second scenario, that the true output variability in each operation is one. For the FSM-All operation, the more likely scenario depends on the estimator used. The value of the FE/IV estimator is more likely to be observed if $\gamma=1$ while the value of the FD/IV estimator is more likely if $\gamma = \gamma_{USPS}$. Finally, the very low point estimates we observe for the FSM1000 operation are much more likely to be observed if the USPS estimates are correct. Overall, our point estimates for manual operations are more consistent with the commission's assumption that the true volume variability equals one than they are with the USPS estimate. For mechanized operations, the answer depends on which estimator is used and how highly disaggregated the FSM category is.

As described in section II.D, an aggregate output variability for flat-sorting operations can be constructed for each plant and time observation using equation (4). This provides a measure of how the total amount of labor in flat-sorting operations is affected by an increase in the number of FHP flats sorted in the plant. The aggregate variability is a share-weighted sum of the output variabilities of the manual, FSM881, and FSM1000 operations. Alternatively, we can ignore the distinction between the 881 and 1000 categories and aggregate over the manual and FSM-all categories. While the log-linear regression specification imposes that the output variability of a sorting operation is the same across all observations, the aggregate variability for flat sorting will vary across observations because the share of hours devoted to each operation varies across plants and time. The mean estimate across all observations, and the standard error of the mean, are reported in the first column of table 7.[35] The top five rows of the table use the disaggregated FSM881 and FSM1000 operations. Using the FE/IV estimator (row 4) the mean estimate across all observations is .838 with a standard error of .046 and for the FD/IV estimator the mean estimate is .914 with a standard error of .029. In both cases these values indicate that the lower

---

[35] The mean variability for flat sorting is constructed as

$$\bar{\varepsilon}_F = \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{3} S_{ij}\hat{\gamma}_j\right) = \frac{1}{n}\sum_{j}\left(\sum_{i} S_{ij}\right)\hat{\gamma}_j$$

where *i* are the plant-time observations, *j* are the 3 MODS operations in flat-sorting, S_{ij} is the share of man hours in flat-sorting that arise in operation *j*, and $\hat{\gamma}_j$ is the estimated output variability for operation *j*. The variance of $\bar{\varepsilon}_F$ is constructed as

$$V(\bar{\varepsilon}_F) = \left(\frac{1}{n^2}\right)\sum_{j}\left(\sum_{i} S_{ij}\right)^2 V(\hat{\gamma}_j).$$

variabilities for the mechanized operations counterbalance the constant returns in manual operations to give an overall variability less than one in flat-sorting operations. This conclusion must be modified slightly if we rely on the FSM-all category. Recall that in Table 4 the output variabilities for this category were generally larger than for the separate FSM881 and FSM1000 operations. As a result, the aggregate variabilities for flat-sorting operations using this category are larger, .925 (.048) and .956 (.029) using the two estimators.

Overall, the two econometric methods and the two different ways of incorporating the mechanized operations yield four estimates of the average output variability for flat sorting. They vary from .838 to .956 with the lower estimates being observed when the mechanized operations are disaggregated by the type of machinery used. Most of this difference in turn stems from the low variability estimates reported in table 4 for the FSM1000 operation. In all cases, however, the aggregate variability is substantially larger than the estimates that do not utilize IV estimators. In particular, the FE estimator advocated in the USPS study yields an variability estimate for flat sorting of .627.

VI.B Output Variabilities in Letter Sorting Operations

The bottom half of Table 4 reports output elasticities of labor demand for the six component operations in letter sorting. Beginning with the OLS estimates in column 1, the output variability estimate for manual operations, .853, appears unreasonably low when compared with the results for manual flat sorting. It is the only letter-sorting operation that indicates significant increasing returns to labor when estimated with OLS.

When we correct for omitted variables bias with either the FE or FD estimator, a familiar pattern appears for all six operations. The output variabilities are substantially smaller than the OLS estimates and indicate significant increasing returns to labor in every operation. Finally, the IV estimates in columns 4 and 5 are all larger than the corresponding non-IV estimates in columns 2 and 3. This basic pattern parallels the differences seen in the flat-sorting operations and again indicates that the endogeneity of output is an issue in the estimation.[36]

Focusing on the FE/IV estimates in column 4, for four of the categories, manual sorting, LSM, OCR, and DBCS, we cannot reject the hypothesis that the output variability is equal to 1.0. Although the OCR point estimate is .882 and the DBCS estimate is 1.241, the standard errors for both estimates are large enough that it is not possible to reject that the estimates equal 1.0.[37] The remaining disaggregated category, BCS (17), has an elasticity of .682, although again the standard error is fairly large. Aggregating together the BCS and DBCS categories, produces an overall BCS estimate of 1.218, which is significantly different than one. Thus, the FE/IV estimates do not provide much evidence for increasing returns to labor in letter-sorting operations. The final column in Table 4 reports FD/IV estimates. One of the estimates (BCS(17)) is .851 and all the other estimates lie in a very narrow range between .972 and .998. We do not reject that the FD/IV estimate is equal to one in each operation. The FD/IV estimates indicate constant returns to labor in all the letter-sorting stages.

The specification tests reported in Table 5, parallel the findings for flat-sorting operations. We always reject the hypothesis that there is no plant effect. In 9 of the 12 cases we also reject the

---

[36] The reduced-forms using FHP flats as the IV produce t-statistics of 9.54 and 10.05 on the FHP flats variable in levels and growth rate regressions, respectively. The FHP flats variable satisfies one of the requirements for a good instrument.

[37] An increase in standard errors relative to OLS is a common occurrence when using IV estimators. The reduction in bias from using the IV estimator comes at the cost of less precision.

hypothesis that output is exogenous. The one operation where we do not reject exogeneity is BCS(17) and this stems from the relatively large standard errors of the coefficients and not from the fact that the IV and non-IV point estimates are very similar. Together the specification tests argue for the inclusion of plant effects and the use of IV estimators.

Table 6 reports the p-values for the hypothesis tests that the true value of the output variability is equal to the USPS estimates and the true value is equal to one. The results are easy to summarize. With the exception of one operation, the p-value for the test that the true output variability equals one is always higher, often substantially higher, than the p-value from the test that the USPS estimates are correct. Alternatively, our point estimates are much more likely to be observed if the true output variability is one than if the true output variability is equal to the USPS estimate. The one exception is with the disaggregated BCS(17) operation. Here the relatively low point estimates we observe (.682 and .851 depending on estimator) are more likely to be observed if the true variability for this operation is .94, the USPS estimate, rather than 1. Overall, the results in Table 6 confirm that our IV estimates for letter sorting operations are more consistent with constant returns to labor than the large degree of increasing returns estimated in the USPS study.

The aggregate output variability for letter sorting is constructed using equation (3) and reported in Table 7. Using the disaggregated BCS/DBCS categories, the average value across all plant observations is 1.026 for the FE/IV estimator and .980 for the FD/IV estimator. We do not reject that each estimate is equal to 1.0 at the .01 significance level. Relying on the single aggregate BCS category, the estimates are 1.025 with the FE/IV estimator and .951 with the FD/IV estimator. Again, the estimates are not significantly different than one, although the last estimate depends on the significance level of the test. Overall, the estimates range from a low of .951 to and high of 1.026 and indicate that constant returns to labor in letter sorting activities appears to be an accurate description of the technology.

As a final summary statistic, we can aggregate over all letter and flat sorting operations to create a plant-level estimate of the output variability using equation (5).[38] This measures the effect on the plant's total labor use in flat and letter sorting that results from an expansion of both the number of flats and number of letters entering the plant. The means over all plant observations are reported in the last column of Table 7. The four estimates, based on the two IV estimates and the two ways of aggregating BCS operations, vary from a low of .952 to a high of .992. Three of the four plant-level means are not significantly different than one. Overall, the plant estimate is a combination of variabilities that are very close to one in most letter-sorting operations and manual flat sorting and variabilites that are less than one in mechanized flat sorting operations. In all cases the plant-level estimates based on the IV estimators are much larger than either the FE or FD estimates that do not fully correct for the endogeneity of output. As a point of comparison, the FE estimator produces a composite plant-level variability of .683, with a standard error of .027, which is very similar to the composite estimate of .71 in the USPS analysis.

VI.C The Effect of Capital, Technology, and Wages on Labor Demands

While the output variabilities are the main focus of interest, it is useful to examine the complete set of parameter estimates in order to better understand the sources of plant-level variation in labor demand. Given our finding that IV estimators are preferred in this setting, we

[38] This plant elasticity ignores parcel sorting operations. The hours shares for flats and letters sum to one.

will focus on the coefficient estimates for the FE/IV estimator. These are reported in Table 8 for flat-sorting operations and Table 9 for letter sorting. Because of the large number of coefficients we will focus on summarizing the capital, technology, and relative wage coefficients.

In flat-sorting an increase in the capital stock of mechanized equipment (KFSM) has a significant negative effect on the demand for manual labor and a significant positive effect on the demand for labor to operate flat-sorting machinery, particularly the FSM1000. This capital acts as a substitute for manual labor and a complement with labor in the FSM1000 operations. The same capital stock does not have a significant effect on the demand for labor in the FSM881 category.[39] The capital stock for bar code readers (KBCR) does not have a significant effect on labor use in any of the categories. This may reflect the fact that the KBCR variable includes all bar code reading equipment in the plant, including that used in letter-sorting operations. Finally, the remaining capital stock in the plant (KOTHER) has a significant positive effect on two of the labor demands. This is likely to reflect an overall effect of plant size on labor use.

The two technology variables, TECH19 and TECH20, have large and significant negative effects on the plant's use of manual labor and, as expected, a significant positive effect on the use of labor in the overall FSM categories. These two technologies act as a substitute for manual sorting. The presence of the FSM1000 technology (TECH20=1) reduces the demand for FSM881 labor, while the presence of FSM881 technology (TECH19=1) increases the demand for FSM1000 labor. The latter effect likely reflects the joint use of these two technologies in many plants and the fact that both substitute for manual sorting.

---

[39] In this case it would be desirable to disaggregate the KFSM capital stock into separate stocks of FSM881 and FSM1000 machinery. We could then include both capital stocks in all three demand equations and estimate the effect of an increase in each type of capital on each category of labor use. Unfortunately, the capital data collected do not allow this disaggregation.

Finally, an increase in the relative wage of mechanized vs manual labor has a significant positive effect on the demand for manual sorting labor and a significant negative effect on the demand for labor in the both the overall FSM and the specific FSM881 category. Both of these effects are expected and indicate that the plant substitutes away from the relatively expensive category of labor. In the final category, FSM1000, an increase in the relative wage of labor in the mechanized category raises the demand for labor which is counter to expectations, although the effect is not statistically significant.

Overall, the capital, technology, and wage coefficients in the flat-sorting demand equations accord with our expectations. They indicate substitution of mechanized operations for manual sorting, as reflected in both the capital and technology dummy variables, and substitution among operations in response to wage differentials

The full set of FE/IV coefficients for the letter-sorting categories are reported in Table 9. The coefficients indicate a very sensible pattern of substitution among the different sorting technologies. The presence of letter-sorting machinery in the plant (TECH02=1) reduces the demand for manual labor as expected. It also has a negative effect on the demand for labor in the OCR, overall BCS, and DBCS categories. The latter effect, in particular, picks up the phase out of the LSM technology and introduction of the DBCS technology over the sample period (see Figure 1). The use of the BCS technology in the plant (TECH17=1) has a significant negative effect on the use of DBCS labor, indicating the two technologies are substitutes. The use of the DBCS technology (TECH18=0) also has a significant effect on the use of OCR and BCS labor. Overall the technology dummies indicate that the use of mechanized LSM substitutes for both manual sorting and later automated technologies. The first generation of automated technologies (BCS and OCR) also act as a substitute with the second generation DBCS technology. The introduction of these different technologies in the plant thus shifts the number of manhours of labor among the different sorting operations and can affect the overall level of labor use in the plant.

The capital stock coefficients indicate a significant role for capital in many of the labor demand equations. In the manual labor demand equation, an increase in LSM capital reduces labor use, indicating that the two operations are substitutes, while an increase in KOCR or KOTHER increases manual labor use. None of the effects are particularly large in magnitude, and, as we saw with the technology variables, an increase in automated sorting capital (KBCS) does not have a significant effect on manual labor use. More substantial capital effects are observed in the other labor demand equations. For example, an increase in a particular kind of capital increases the demand for the complementary labor input, as expected, in the LSM, OCR, BCR-All,and DBCS demand equations. The exception to this pattern is that an increase in KBCS lowers the demand for labor in the disaggregated BCS category, although it increases the use of total use of labor in the disaggregated BCS/DBCS categories. This may result from the fact that our BCS capital stock is an aggregate of both BCS and DBCS equipment and its variability is driven by differences in the stock of DBCS equipment. A reasonable pattern of substitution is observed in the LSM demand equation. Here an increase in the capital stock of automated operations (KOCR, KBCR, and KBCS) all lead to a decline in the use of labor in category 02. An increase in the KOCR capital stock leads to an increase in the use of labor in the OCR and BCS operations. All of these patterns suggest that the capital in the automated operations reduces the demand for labor to operate letter-sorting machinery but increases the demand for labor in the automated-sorting operations. This reflects the expected pattern of substitution among the alternative sorting technologies.

The relative wage coefficients indicate the expected pattern of input substitution. An increase in the relative wage of labor in the mechanized sorting operations (WLSM/WMAN) increases the demand for labor in manual and automated (BCS-All and DBCS) categories and reduces the demand in the LSM operation. An increase in the relative wage in automated operations (WAUT/WMAN) raises the demand in the substitute manual and LSM categories while lowering the demand in the automated operations OCR, BCS, and DBCS.

Overall, the coefficient estimates on the technology, capital, and wage rate variables are often statistically significant and indicate a reasonable pattern of substitution among alternative technologies and between high and low-wage categories of labor. In particular, the estimates show that it is important to model the labor demand curve for a particular operation as part of an integrated set of labor demand equations that recognize the various technologies and types of capital in place in the plant.

VII. Generalizations of the Labor Demand Models

The labor demand estimates reported in Tables 3, 4, 8, and 9 are based on log-linear models that include the basic set of variables that theory indicates should be important in labor demand equations. The model deals with what we believe are the issues of first-order importance: the specification of the set of explanatory variables and the treatment of the endogenous output variable. It is possible to refine the model further. We have already discussed the desirability of further disaggregation of the FSM and BCS capital stocks, and the need to correct for price changes and depreciation in construction of the capital stocks. In the remainder of this section we will address several additional specification issues that were discussed in the R2000 rate case and assess their importance to the estimates reported here.

VII.A Quadratic Output Effects

The log-linear regression equation restricts the output variability for a sorting operation to be constant across all observations regardless of the time period or size of the plant. It is simple to generalize the model to allow the estimated variability to depend on observable plant characteristics including the output level and type of capital in the plant. All that is necessary is to include powers of q and interactions between the relevant observable characteristics and output (Xq) as additional explanatory variables.[40] This technique was used by the Postal Service to specify translog models of labor demand. This extension of the model is useful if there is interest in how, or whether, the output variability changes across the set of plants or over time. One potential difficulty with the extension is that the additional interaction terms may be fairly highly correlated with the initial output variables and this will result in larger standard errors for all the output coefficients.

In order to assess whether the output variability differs between small and large plants, we generalize the model in the simplest possible way by including the square of output as an additional explanatory variable. The empirical model is now:

$$(19) \qquad h_{it} = \beta X_{it} + \gamma_1 q_{it} + \gamma_2 (q_{it})^2 + m_i + \varepsilon_{it}$$

[40] When output is endogenous these additional interaction terms will also be endogenous and at least one unique instrumental variable will be necessary for each of the new explanatory variables. Given an instrument z, this can be handled by using powers of z and interactions between X and z as additional instruments.

The estimated output variability is now measured as $\gamma_1 + 2\gamma_2 q_{it}$ and will vary with the level of output in the plant. The output coefficient estimates for this quadratic model are reported in Table 10 for both the FD/IV and FE/IV estimator. The first column reports the estimate of γ_1 and the second column reports the coefficient estimate of γ_2. With the exception of the FSM1000 , one of the disaggregated BCS estimates, and one of the manual estimates, the estimate of γ_1 is always statistically significant. In contrast, of the 20 estimated values of γ_2, only one is statistically significant. The implication is that, with the exception of the FE/IV estimate for manual letter sorting, all changes in the output variabilities between plants of different sizes will be due to statistically insignificant coefficients.

Focusing on the FD/IV estimates in the top half of the table, we observe that the estimate of γ_1 is very similar to the estimated output variability in the last column of Table 4. The estimate of γ_2, in addition to being statistically insignificant, is generally very small and this leads to trivial variation in the output variabilities across plants of different sizes. The exception to this pattern are the estimate for the FSM1000 operation and the LSM category. The mean estimate of the variability across all observations is reported in the third column and, again with the exception of these two groups, are all close to or above 1.0. The final column of Table 9 reports the value of the output elasticity at the 25th and 75th percentile of the plant output distribution. The first number is the variability estimate for a plant with the output level of the 25th percentile across all observations. The second number is for a plant with the output level equal to the 75th percentile of the output distribution. With the exception of the FSM1000 and LSM categories, which have large but statistically insignificant estimates of γ_2, the numbers in the final column indicate an extremely narrow range of estimates. For example, in manual flat sorting operation, plants in the middle 50% of the output distribution have estimated variabilities that lie between .969 and .973. Overall, the FD/IV estimates indicate that there is no value to generalizing the log-linear model in this way.

The estimates using the FE/IV estimator are reported in the bottom half of the table. In this case the estimates of γ_1 are not identical to the estimates in Table 4 and the estimates of γ_2 are generally larger than for the the FD/IV estimator. The mean variability estimate across all plants, however, is fairly similar to the linear model estimates for the FE/IV estimator in Table 4. The larger estimate of γ_2 leads to more variation in the output variability across plants, as evidenced by a wider range of estimates in column 4, but, with the exception of the one manual operation, all the across-plant differences are generated by an insignificant γ_2 coefficient. The one exception is the manual letter sorting category. Here the quadratic coefficient is statistically significant and positive, which implies that the output variability rises with plant size. The middle half of the estimates vary between 1.014 to 1.185 indicating decreasing returns to manual sorting for many of the observations. Overall, given the low significance levels on the coefficients for q^2, there is little reason to prefer the quadratic estimates to the linear model estimates reported in Tables 4.

VII.B Characteristics of the Plant's Service Area

A second way to generalize the model is to allow characteristics of the plant's service area to shift the labor demand equations. In the analysis conducted by the Postal Service in USPS-T-14, they include the log of the number of delivery points (DPT) in the plant's service area as an additional shifter in the labor demand equations. As explained in section III.F, it is possible to rationalize including this variable. We have estimated the demand equations with this variable included and the effect on the output variability is summarized in Table 11. The results are easy to summarize. The DPT variable is statistically significant in less than half of the operations and, more importantly, has no appreciable impact on the estimated output variabilities.

The first two columns report the coefficient on the DPT variable for the two IV estimators. There is actually very little sample variation in this variable. It tends to increase slowly over time but does not differ dramatically across plants. When it is included in the FD specification it has a small but statistically significant coefficient in three sorting operations. When it is included in the FE specification it acts very much like the constant term in the regression. The magnitude of the coefficient is large but the constant term in the regression changes to offset it, and the variable contributes little to the explanatory power of the model. Most importantly, the inclusion of the DPT variable has no effect on the estimated output variabilities. Comparing the estimates in the last two columns of Table 11 with the estimates in the last two columns of Table 4, it is clear that controlling for the number of delivery points in the plant's service area has no impact on inferences about the output variabilities in our framework.

VII.C Lagged Adjustment to Output Changes

An additional specification issue we explore involves the use of lagged values of the output variable in the labor demand equation. One rationale for including lags is that there is slow adjustment in the mix of sorting operations used within the plant to changes in output. The demand equations estimated in the Postal Service analysis included output measures defined at the level of each sorting operation and showed evidence of slow adjustment. To examine this issue, we add three lags of the output variable to each labor demand equation so the estimating equation becomes:

$$(15) \qquad h_{it} = \beta X_{it} + \gamma_0\, q_{it} + \gamma_1\, q_{it-1} + \gamma_2\, q_{it-2} + \gamma_3\, q_{it-3} + m_i + \varepsilon_{it}$$

The output variability is defined as the proportional change in h given a proportional change in all output variables and is measured as the sum of the output coefficients $\gamma = \gamma_0 + \gamma_1 + \gamma_2 + \gamma_3.$ When using the IV estimator, three lags of the instrument are also included as additional instruments. When using the FD estimator there will be four output growth rates, one for the current period and three lags, as explanatory variables.

The coefficient estimates for the four output coefficients are reported in Table 12. The FD/IV estimates are reported in the top half of the table. The results are simple to summarize. The current period output change is a significant determinant for all but one operation and the lagged growth rates are virtually never significant. Only 6 of the 30 lagged coefficients are significant at the .01 level and half of those are for the aggregated BCS-All category. The overall output variability, defined as the sum of the coefficients, is slightly larger than the FD/IV estimate reported in Table 4 for six operations (05, 20, 06, 02, 10, 18) and smaller for the other three operations (19,01,17). Because the output variability for an operation is now the sum of four coefficients, three of which are usually statistically insignificant, it has a standard error that is much larger, generally two to three times, than the single coefficient estimates in Table 4. The only case where we reject the hypothesis that $\gamma = 1$ is for the DBCS operation. In that case the point estimate is 1.901. Overall, adding lags in the FD model simply introduces a great deal of imprecision into the estimated output variabilities. This does not appear to be an improvement over the original estimates reported in Table 4.

The FE/IV estimates are reported in the bottom of Table 12. We observe a similar pattern to the FD/IV estimates. The current output level is a significant variable, but the three lagged values of output are statistically significant in only 3 of 30 cases. Sometimes the magnitude of the lagged coefficients can be fairly large, as much as one-third of the magnitude of the coefficient on current output, but they are never precisely estimated. The large number of insignificant coefficients generates a large standard error on the output variability reported in the last column. For four of the operations, 05, 20, 02, and 17, the output variability from Table 12 is larger than estimate in Table 4, although the estimates are not very precise. For the other operations, the estimates to the coefficients in Table 4With the exception of the FSM1000 operation, the point estimates for the output variabilities are fairly similar to the estimates reported in Table 4, but the standard errors are as much as 25% larger than the standard errors in Table 4. Of the 10 estimates using the FE/IV estimator in Table 12, only two, for operations 19 and 10, are significantly different than one. The conclusion with the FE estimator is that including lagged values of the output variable has little effect on the estimated output variability for most of the operations but decreases the precision of the estimates. This model does not improve the estimates relative to the simple model with a single output variable and no lags.

VII.D The Use of TPF as the Output Variable

As described in Sections II and III, there are several specification differences between the model estimated by the Postal Service and the model we use in this paper. Some of the differences are the result of modeling choices that must be made in an empirical project. These include the functional form of the regression equation, the correction for characteristics of the plant's service area, and the specification of the lag structure for labor use. Each of these assumptions has been examined in Sections VII.A. - VII.C. and found to not be important in determining the results.

A second set of differences between this paper and the earlier work follow directly from the model of production used here. These include the use of disaggregated capital stocks in both manual and automated sorting operations, the use of relative wages, the elimination of the "manual ratio" in the manual sorting equations, and the specification of FHP counts for letters and flats as the appropriate output variables. In each of these cases the theoretical model developed in Section II provides guidance on the form of the estimating equations and these aspects of the empirical model cannot be altered without respecifying the underlying model of production developed there.

The model of production developed here views the letter-sorting process as using five types of labor input (hours in OCR, LSM, Manual, BCS and DBCS operations) and five types of capital input (capital in OCR, LSM, BCS, BCR, and other) to sort a number of pieces of mail which is measured by the FHP count for letters. Similarly, the flat-sorting process combines 3 types of labor and 3 types of capital to sort pieces of mail measured by the FHP count of flats.

The alternative model presented by the Postal Service uses a different measure of output in each sorting operation. The output is a measure of the number of pieces of mail fed through the machinery or otherwise handled in the specific operation (TPF by operation).[41] The underlying form of the production function for mail sorting that would result in the use of TPF by operation as the output variable is not specified in the Postal Service analysis. Nevertheless, in order to try to

---

[41] The output variable they use in a mechanized or automated sorting operation is the machine count of pieces fed (TPF) in that operation. In manual operations they use the total number of pieces handled (TPH) in the operation as the output variable. See R2000-1-T-15, p. 50 for discussion. We will refer to both as TPF.

isolate the sources of difference in the estimated output variabilities between this paper and the Postal Service analysis, we reestimate our model using TPF in an operation as the output variable in that operation. That is, we replicate the steps underlying the estimates in Table 4, making only a change in the output variable used in the equation. In each equation, FHP for the shape will be replaced with the Postal Service's measure of output in the operation.

The econometric issues discussed in Section III are still relevant when this change is made. The output variable is likely to be endogenous for all three reasons discussed in Section III. B: omitted variables, measurement errors, and the simultaneous choice of hours and output in an operation.[42] We will use First-Difference and Fixed Effects Estimators to correct for omitted plant-specific variables and use the IV method to correct for the endogeneity arising from output measurement error and simultaneity. As with the estimates reported in Tables 3 - 9 we will use the FHP count of letters as the instrument for TPF in each of the three flat-sorting operations. In these equations, the FHP counts of letters satisfies the requirements of an instrument. It is likely to be correlated with the TPF count in the operation but uncorrelated with the error term in the equation. In particular, the error term is composed of measurement errors in output and random shocks to hours that also lead to changes in the TPF count in the operation. That is, the count of FHP letters in the plant is exogenous in the plant's choice of hours and piece handlings in each flat sorting operation. For the same reasons we will continue to use the FHP count of flats as the instrument for the TPF count in each letter-sorting operation.

In the model we developed in Section III of this paper we did not estimate labor demand

[42] As described above, only the first two sources of endogeneity are likely to be relevant for our specification using FHP as the output variable.

equations for parcel or priority sorting operations because we could not develop a measure of FHP for the parcel operations and the fact that MODS operation 12, the small parcel and bundle sorter, handles both priority and parcel pieces. By using TPF as the output variable for an operation, the Postal Service analysis is able to estimate labor demand equations for three additional sorting operations: manual priority sorting (MODS 07), manual parcel sorting (MODS 08) and small parcel bundle sorter, SPBS (MODS 12). We will also estimate these three factor demands using TPF as the output measure. In all three equations we will use both FHP letters and FHP flats as instrumental variables for TPF in the equation.[43]

Table 13 reports the output coefficients for the 13 sorting operations. It is the same form as Table 4 and the underlying model is identical except for the output variable. Comparing the alternative estimators we see a familiar pattern: first-difference and fixed effects estimates in columns 2 and 3 are less than the OLS estimates in column 1. This is true for every sorting operation. This pattern was also seen when comparing OLS and FE estimates in the Postal Service analysis USPS-T-14. The FE estimates across the 13 operations vary from a low of .216 (in manual parcel sorting) to a high of .871 (in LSM sorting). As with the Postal Service analysis, the manual operations tend to have low elasticities. The four manual sorting operations have variability estimates that tend to be at the low end of the 13 categories.

The assumption that output is exogenous in these equations can be tested with the Hausman test. While we will not report a table of the results, we reject the exogeneity of output in

---

[43] The other variables included in the labor demand equations for mods operations 07, 08, and 12 follow the general specification in section II. The variables are the capital stock in the SPBS operation (KSPBS), the capital stock in non-sorting operations (KOTHER), the relative wage of automated to manual labor, a technology dummy indicating if the plant used the SPBS technology (included in operations 07 and 08), a time trend, quarterly dummies, and a dummy variable for the change in FHP conversion factors .

in 9 of the 13 categories using the FE/IV estimator and 9 of the 13 categories using the FD/IV estimator.[44] IV estimators that control for the endogeneity of output are reported in the last two columns of Table 13. In every one of the 26 cases, the correction for endogeneity results in an increase, and often a substantial increase, in the output variabilities. Of the 26 IV estimates, all but six are greater than .90 and 21 of the estimates are not significantly different than one. Based on the IV estimates in the last two columns of Table 13, constant or diminishing returns to labor would be the norm in virtually all letter and flat sorting operations. This contrasts strongly with the estimates based on the FE estimator in column 2, which indicate increasing returns to labor in all sorting categories.

We can also compare the IV estimates in Table 12 that use TPF as the output definition with the IV estimates from Table 4 that use FHP as the output variable. Here the differences across output definitions are much less substantial than the differences between IV and non-IV estimators. The use of TPF as the output variable results in slightly larger estimates of the output variability than the use of FHP in 8 of the 10 letter and flat sorting categories when using the FD/IV estimator. In fact, all of the automated and mechanized operations have variabilities greater than one when TPF is used as the output definition. When using the FE/IV estimator the estimated variabilities in Table 13 are larger than their FHP counterparts in Table 4 for 6 of the 10 letter and flat sorting categories, but, again, the differences are not as large as the differences across estimators.

[44] For three operations, FSM1000, LSM, and Manual Parcels, we do not reject the exogeneity of output using either estimator. While the IV estimates are always larger than the corresponding non-IV estimate, the IV estimates often have large standard errors which leads to the failure to reject exogeneity. In the other two cases where we fail to reject exogeneity of output, the FE and FE/IV estimates for the disaggregated BCS category are very similar, while the FD and FD/IV estimates for the SPBS category are different but, as seen in most of the other cases, the IV estimate is not very precise.

The important conclusion from Table 13 is that the correction for endogeneity is crucial, regardless of whether output is measured with FHP or TPF. Estimators such as the FE or FD estimators that fail to account for the endogeneity of output underestimate the output variabilities. While we do not agree with the use of TPF as the output variable in the labor demand equations and, thus, do not rely on the estimates in Table 13, it is the case that the treatment of endogeneity is of major importance in the estimation of labor demand equations for the sorting operations.

VII.E Labor Demands for Priority Mail Sorting

In the model developed in Section III we focused on the ten operations used to sort letters and flats. In this section we extend the framework to include two sorting operations used for priority mail. [45] We will develop and estimate labor demand equations for manual sorting of priority mail (MODS category 08) and sorting of priority mail on the small parcel and bundle sorter (MODS 04). The latter category was not analyzed separately by the Postal Service in either R2000 or R2001, but was aggregated with parcels that are sorted on the SPBS to create a labor demand equation representing the total use of labor on the SPBS machinery (MODS 12).

Using the framework developed in Section II, we treat priority mail as separable from letter and flat sorting. In fact, simply labeling the third category of mail as priority rather than parcels, allows us to use the entire setup from Section II to specify the demand equations for priority mail sorting operations. We can measure the number of FHP priority pieces of mail as the sum of the FHP count in MODS operations 08 and 04. This is the output variable we will use for the labor demand equations. In the estimating equations we will also include the capital stock in the SPBS

---

[45] As discussed in footnote 6, it is not possible to construct a measure of the FHP count for parcels and this makes it impossible for us to estimate labor demand equations for parcel-sorting operations.

operation (KSPBS), the capital stock in non-sorting operations (KOTHER), the relative wage of automated to manual labor, a technology dummy indicating if the plant used the SPBS technology (included in the demand equations for the manual operation 07), a time trend, quarterly dummies, and a dummy variable for the change in FHP conversion factors

There is one significant difference between priority and letter or flat sorting operations. While letter and flat sorting is conducted in all plants and a measure of FHP flats and FHP letters is available for virtually all plant-quarter observations, priority mail is not sorted in all plants or in all time periods. Of the 301 plants we utilize in this study, between 222 and 257 report a positive value of FHP for priority mail in any quarter. In addition, the number of plants that sort priority mail trends downward over time and the decline is particularly noticeable after 1997. This appears to reflect a consolidation of priority sorting operations in a smaller number of plants. Of the 301 plants we utilize, 158 sort priority mail in all of the 28 quarters we examine, 16 never sort it in any quarter, and the remaining 127 plants sort it for a subset of the time periods. This raises the question of whether the choice to sort priority mail is related to the plant's size. If so, it is necessary to be more specific in modeling the process that relates plant size, the choice to do priority sorting, and the hours of labor utilized. This is beyond the scope of the current study. As a simpler alternative we will estimate the relationship between output and labor use in priority mail sorting for the subset of 158 plants that always engage in this operation. As a result our inferences will be limited to the output-labor adjustment process in the set of plants that have a long-term commitment to this type of mail sorting. While these plants may represent a non-random subset of all the plants that sort priority mail, they do account for the vast majority of the priority sorting that is done by the USPS. Over the period of our data, the share of all FHP priority mail sorted in these 158 plants rose from 67% to 86% over time.

The estimates of the output variabilities using the different estimators are reported in Table 14. The same general pattern seen in Tables 4 and 13 is evident. FE and FD estimates are substantially lower than OLS and IV estimates. The Hausman test indicates that we reject the exogeneity of output in the manual equation using both estimators, and we also reject the exogeneity of output in the SPBS labor demand using the FE estimator. Again, the evidence indicates the use of IV methods is necessary.[46]

Focusing directly on the IV estimates we see that the two methods produce somewhat different estimates. The FE/IV estimates are 1.105 and 1.898 for the manual and automated operations, respectively. Neither estimate is significantly different than 1, although, in the SPBS operation, it is because the standard error is very large. The FD/IV estimates are both smaller. For manual sorting the variability is .893 and for sorting on the SPBS it is 1.054. We do not reject that either estimate is equal to 1. Overall, our preliminary evaluation of the two priority mail sorting operations allows two broad conclusions. First, as with the other operations, the treatment of the endogeneity of output makes a substantial difference to the estimates. Second, once the endogeneity of output is addressed with IV estimators, there is no evidence of substantial increasing returns to labor in either the manual or mechanized sorting operations.

VIII. Summary and Conclusion

---

[46] The reduced form regressions indicate that the count of FHP letters is significantly correlated with the FHP of priority mail but the FHP count of flats is not. FHP letters is the better instrumental variable to use in correcting for the endogeneity of output in the priority sorting operations.

This paper reexamines the estimation of the output variability of labor demand using plant-level panel data for mail sorting facilities. The basic approach, introduced into the R2000 rate case in USPS-T-15 and expanded in R2001 USPS-T-14, of modeling labor demand by MODS category using plant-level panel data is appropriate for estimating the output variabilities for different sorting operations. In this paper we utilize the data analyzed in USPS-T-14 but make several changes to the empirical labor demand equation and econometric methods that are used. Specifically, the model used here views each plant as producing two outputs, a number of sorted letters and a number of sorted flats, using eight different types of labor inputs and six types of capital equipment. Labor demand equations are specified for each of the eight categories of labor and are a function of the output of FHP letters or flats sorted in the plant, the capital stocks, relative wages of different types of labor, and dummy variables measuring the mix of sorting technologies used in the plant.

The econometric methods that are used to estimate the labor demand equations specifically focus on the endogeneity of output in the labor demand equations and the biases this is likely to induce in ordinary least squares estimates. We show that endogeneity is likely to arise from two sources: omitted variables that are correlated with the plant's output level and measurement error in output. We propose two estimators based on the method of instrumental variables that can address both of the sources of endogeneity. The estimators combine the use of plant-specific intercepts, to deal with omitted variables problems, and instrumental variables, to address measurement error problems. Specification tests indicate that the assumption that output is exogenous is rejected in most demand equations and that the econometric methods we adopt are necessary to produce consistent estimates of the output variabilities.

The output variability estimates from the instrumental variables (IV) estimators are reported in the last two columns of Table 4. In the case of the six operations involved in letter sorting, the Fixed Effects/IV estimator produces estimates of the output variabilities between .682 and 1.241 depending on operation. In four cases, labor used in LSM, OCR, and Manual operations, we do not reject the hypothesis that they are equal to one. One operation, BCS sorting, shows evidence of increasing returns to labor with a point estimate of .682. When we examine the same operations using the First Difference/IV estimator we find that one of the estimates is .851 and the other five all lie between .972 and .998. None of the six estimates are statistically different than one. Aggregating over these sorting operations to produce an output variability for overall letter-sorting operations, we get a mean estimate across all sample observations that varies from .951 to 1.026 depending on the econometric estimator used and the level of disaggregation of the BCS operations.

Focusing on flat-sorting operations, we estimate the output variability for manual sorting to be .884 and .961 with the two IV estimators. If we combine all flat sorting machinery operations into a single FSM aggregate, we get output variability estimates of .963 and .916 depending on the estimator used. Neither estimate is significantly different than one. If we disaggregate FSM operations into separate categories based on the type of equipment used, we find some evidence that output variabilities are less than one, particularly for the FSM1000. Aggregating over the flat-sorting operations we get estimates of the average output variability for flat sorting that lie between .838 and .956 depending on estimator and the level of disaggregation of FSM operations.

We also produce a single estimate of the output variability for each plant as a weighted sum of the variabilities in all the letter and flat sorting categories. This estimate differs across plants because of differences in the intensity of use of the different sorting operations. The estimates of the aggregate plant-level output variability, averaged over all plant-time observations, varies from .952 to .992 depending on the estimator used and the degree of disaggregation in the underlying sorting operations. These estimates are substantially larger than the composite estimate for all plant operations of .71 reported in USPS-T-14.

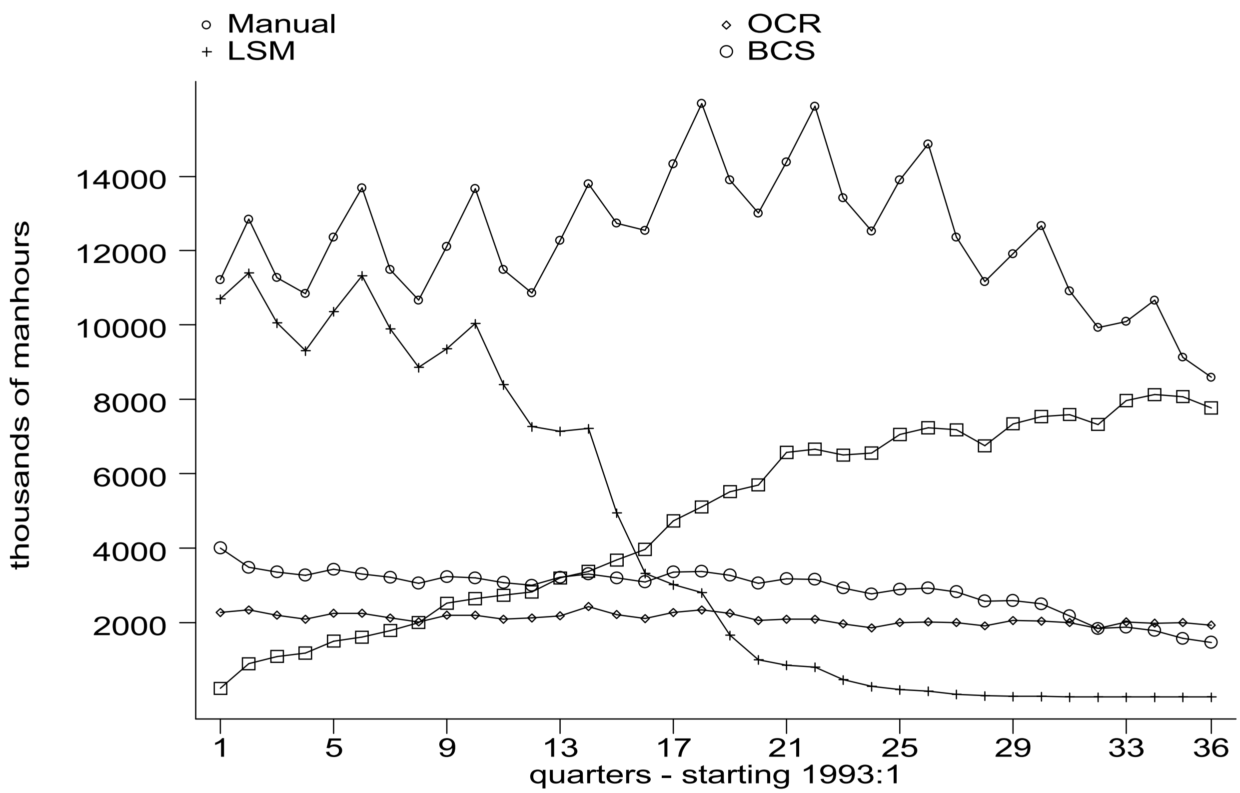
In general, the estimates of the output variabilities using the IV estimators developed in this paper are larger and closer to one than the estimates produced in the Postal Service analysis in either R2000 or R2001. The estimates here imply that we cannot reject that there are constant returns to labor in many of the sorting operations. This is particularly true in the manual sorting operations for both letters and flats. While there are numerous differences in model specification between this study and the Postal Service analysis, the main source of difference lies in the treatment of the endogeneity of output in the labor demand equations. We find that output is endogenous in the labor demand equations and use instrumental variable estimators to control for this problem. This correction results in a significant increase in the magnitude of the estimated output variabilities. The model used in this paper is based on a different definition of output than the Postal Service analysis in USPS-T-14. Even when we adopt the output definition from the USPS study, we find that we reject that output is exogenous in virtually all operations and the use of IV estimators results in substantially larger estimates of the output variabilities.

Other differences in the results reflect the differences in specification and measurement of several key variables. We find an important role for specific types of capital in many of the demand equations and reasonable substitution effects in response to wage differences. We do not find that generalizations of the model that (1) include a quadratic term in output (so that the output variability differs across plants of different sizes), (2) include the number of delivery points in the plant's service area or (3) allow for slow adjustment of the labor input over time have any significant effect on the estimated output variabilities. At best, they have no effect and, at worst, they reduce the precision of the estimates.

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)

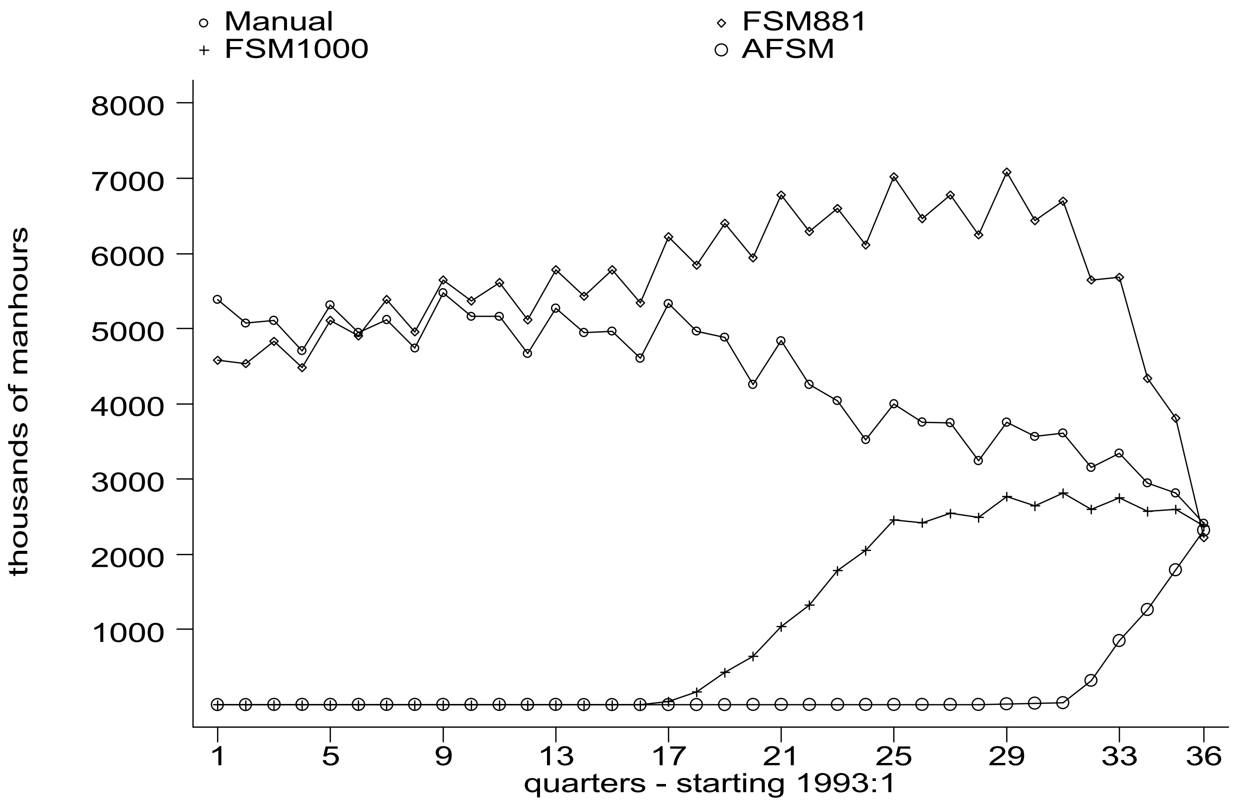


Figure 3

Effect of Change in Conversion Factors on Aggregate FHP Letters

(Sum over 321 plants)

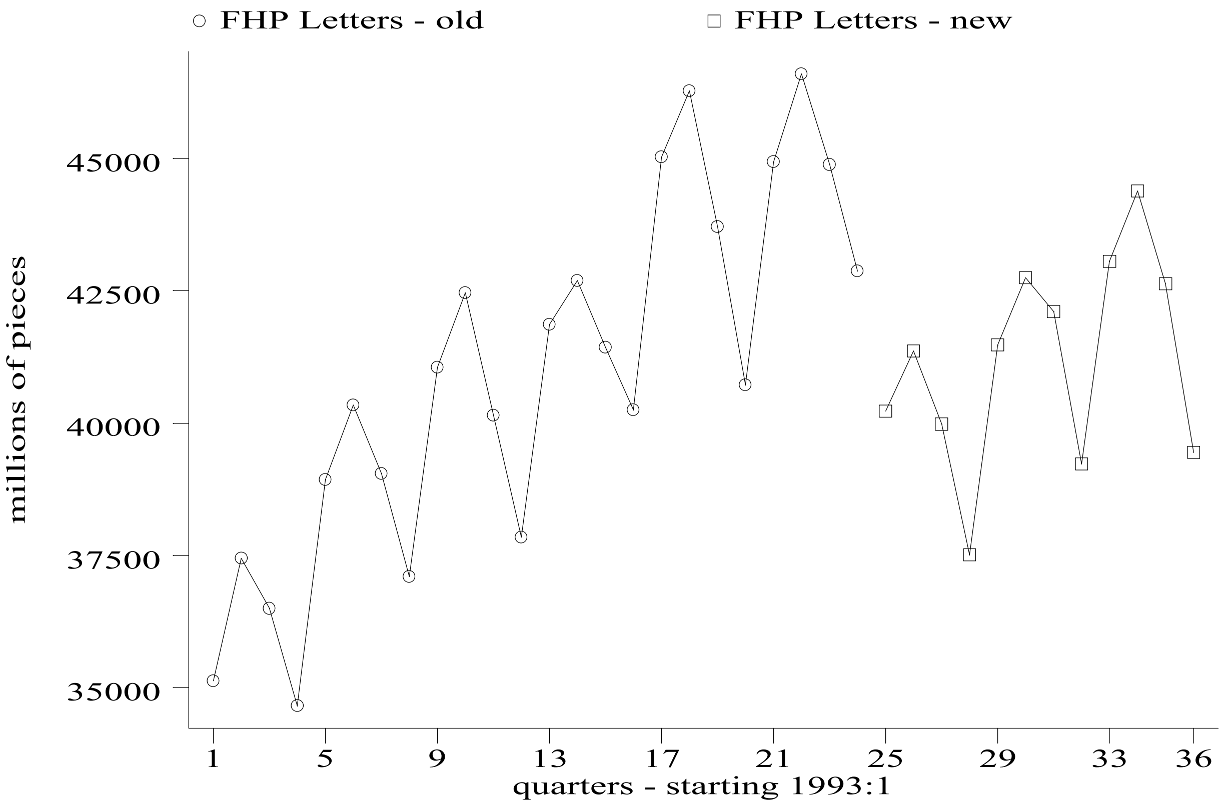


Figure 4

Effect of Change in Conversion Factors on Aggregate FHP Flats

(Sum over 321 plants)

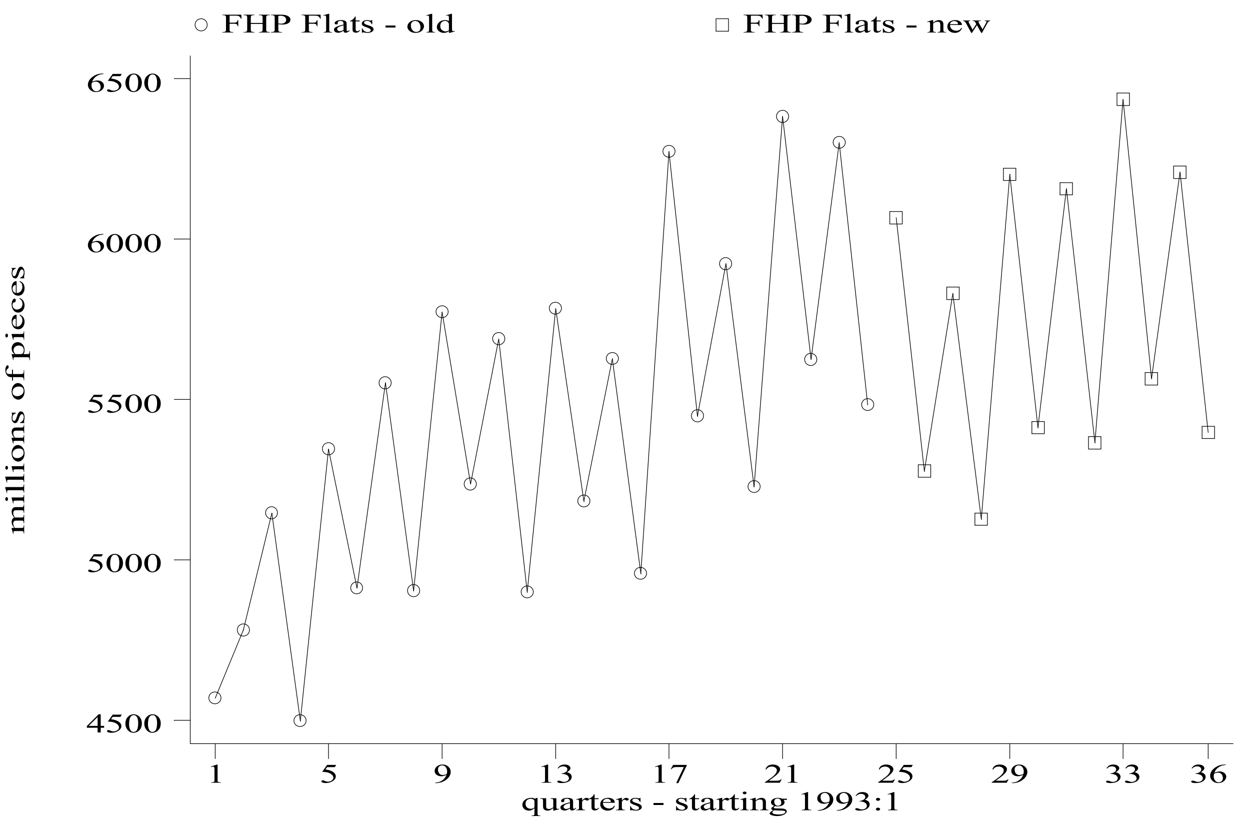


Table 1

Comparison of Model Specification and Econometric Methods with R2001-1, USPS-T-14

	This study	R2001-1 USPS-T-14
Mods Sorting Operations Covered	Letters- 06, 01, 02, 17, 18 Flats - 05, 19, 20 Priority -04, 08 in section VII.E	Letters - 06, 01, 02, 17, 18 Flats - 05, 19, 20 Parcels/Priority - 07, 08, 12
Output Variable	FHP by shape (letters/flats/priority) Sensitivity test to TPF/TPH in section VII.D	TPF by operation for automated operations, TPH by operation for manual operations
Lagged Output	Not included in base model. Sensitivity test in section VII.C.	Four lags included
Capital Variable	Six categories of assets. Included in automated and manual operations.	Total plant capital (automated letter sorting capital in two operations). Not included in manual operations.
Wage Variable	Relative wage of automated/manual workers	Wage rate in the operation
Cyclical and Trend Variables	Quarterly dummies, Trend, Conversion factor on intercept	Quarterly dummies, Trend, Conversion factor on slopes
Service Area Variables- Number of Delivery Points	Not included in base model. Sensitivity test in section VII.B	Included
Other Variables	Dummies for other automated/mechanized technologies used	"Manual Ratio" in manual operations
Functional Form	Linear. Sensitivity test for quadratic in output in section VII.A	Quadratic in all variables
Econometric Methods	Fixed Effects Estimator, First Difference Estimator, Hausman Test for Endogeneity, IV Estimators	Fixed Effects Estimator, Fixed Effects Estimator with AR1 correction.

Table 2

Technology Use

(Proportion of 321 plants that use a technology)

Year: quarter	93:1	94:1	95:1	96:1	97:1	98:1	99:1	00:1	01:1
Letter Sorting									
LSM (02)	.885	.897	.897	.879	.723	.258	.087	.022	.006
OCR (01)	.897	.922	.931	.938	.935	.851	.847	.854	.857
BCS (17)	.922	.922	.907	.891	.888	.822	.816	.832	.829
DBCS (18)	.417	.701	.826	.841	.910	.963	.966	.950	.953
Manual (06)	.956	.975	.981	.984	.981	.968	.978	.950	.953
Flat Sorting									
FSM 881 (19)	.723	.754	.760	.754	.751	.754	.754	.763	.801
FSM 1000 (20)	0	0	0	0	.044	.380	.607	.620	.626
AFSM	0	0	0	0	0	0	0	0	.318
Manual (05)	.953	.972	.981	.984	.981	.972	.972	.944	.938

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

Table 4

Output Variability of Labor Demand: Alternative Estimators

(robust standard errors in parentheses)

	OLS	Fixed Effects	First Difference	Instrumental Variables Fixed Effects	Instrumental Variables First Difference
Flat Sorting Operations					
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 5

Specification Tests

	No plant effect	Hausman test: *Cov (q, ε) = 0* (with robust standard errors)	
	F - stat	FD t-stat	FE t-stat
Flat Sorting Operations			
Manual (05)	32.56*	-8.29*	-3.18*
All FSM (11)	26.02*	-5.17*	- 4.35*
FSM 881 (19)	53.97*	-5.36*	- 5.18*
FSM 1000 (20)	35.81*	-0.26	-1.48
Letter Sorting Operations			
Manual (06)	80.63*	-7.20*	-7.64*
LSM (02)	30.72*	-3.02*	-4.00*
OCR (01)	33.69*	-4.04*	-2.07
All BCS (10)	31.69*	-3.28*	-3.81*
BCS (17)	35.26*	-1.97	-0.86
DBCS (18)	23.97*	-2.88*	-3.61*

* Reject the hypothesis at the .01significance level.

Table 6

Hypothesis Tests Against USPS Estimates and Commission Assumptions

	Point Estimate from R2001-1 USPS-T-14 *	P-Value for Hypothesis Test			
		Null Hypothesis Ho: $\gamma = \gamma_{USPS}$ Alt. Hypothesis Ha: $\gamma > \gamma_{USPS}$		Null Hypothesis Ho: $\gamma = 1$ Alt. Hypothesis Ha: $\gamma < 1$	
		FE/IV	FD/IV	FE/IV	FD/IV
Flat-Sorting Operations					
Manual (05)	.71	.010	<.001	.061	.187
FSM-All (11)	.84	.022	.035	.271	.023
FSM881 (19)	.74	.123	<.001	.348	.097
FSM1000 (20)	.74	.508	.996	<.001	<.001
Letter- Sorting Operations					
Manual (06)	.58	<.001	<.001	.516	.456
LSM (02)	.90	.104	.268	.767	.480
OCR (01)	.77	.092	.001	.079	.520
BCS-All (10)	.93	<.001	.117	.001	.330
BCS-other (17)	.94	.948	.829	.022	.056
DBCS (18)	.87	.011	.051	.067	.488

* Source for FSM-All and BCS-All categories is R2001-1, USPS-T-14, Table 5, Col. 2, p. 51. Source for all other categories is USPS-T-14, Table 6, Col. 1, p. 53

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.

Table 8

Labor Demand Coefficients: Flat Sorting Operations
FE/IV estimator (robust standard errors in parentheses)

	Manual	FSM-All	FSM881	FSM1000
ln (F)	.884(.075)*	.963 (.061)*	.803 (.054)*	.739 (.247)*
Capitol:				
KFSM	-.151 (.023)*	.036 (.007)*	-.004 (.008)	.175 (.029)*
KBCR	.109 (.069)	.034 (.037)	.043 (.034)	-.133 (.082)
KOTHER	.006 (.003)*	-.002 (.001)	.005 (.001)*	-.003 (.004)
Technology:				
TECH19	-.524 (.107)*	1.154 (.413)*	n.a.	.235 (.096)*
TECH20	-.485 (.019)*	.276 (.013)*	-.044 (.010)*	n.a.
Relative wage: $\frac{WFSM}{WMAN}$	.097(.026)*	-.183 (.060)*	-.190 (.159)*	.266 (.137)
DQ2	.053 (.016)*	.055 (.011)*	.031 (.011)	.033 (.040)
DQ3	-.000 (.013)	.006 (.010)	-.011 (.009)	.017 (.022)
DQ4	-.007 (.019)	.050 (.014)*	-.005 (.013)	.060 (.051)
TREND	-.006 (.001)*	.003 (.001)*	.005 (.001)*	.021 (.003)*
DC	-.076 (.019)	.134 (.016)*	-.008 (.012)	.141 (.040)*
$\hat{\sigma}$	.403	.267	.249	.210
R^2	.863	.922	.929	.920
Sample size	7277	6408	6343	1562

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

Table 9

Labor Demand Coefficients: Letter Sorting Operations

FE/IV estimator (robust standard errors in parentheses)

	Manual	LSM	OCR	BCS-All	BCS (17)	DBCS(18)
ln (L)	1.002 (.051)*	1.137 (.188)*	.882 (.084)*	1.219(.057)*	.682 (.158)*	1.241 (.161)*
Capital:						
KOCR	.009 (.003)*	-.135 (.009)*	.040 (.005)*	-.003 (.003)	.059 (.009)*	.028 (.005)*
KBCR	-.024 (.031)	-.474 (.129)*	.000 (.042)	.006 (.032)	.119 (.059)	-.135 (.061)
KBCS	-.001 (.001)	-.044 (.009)*	.003 (.003)	.014 (.002)*	-.024 (.006)*	.044 (.005)*
KLSM	-.035 (.006)*	.246 (.041)*	-.016 (.009)	-.013 (.009)	-.031 (.019)	.027 (.011)*
KOTHER	.007 (.001)*	.001 (.004)	.007 (.002)*	-.010 (.001)*	-.004 (.003)	-.010 (.003)*
Technology:						
TECH02	-.141 (.009)*	n.a.	-.042 (.016)*	-.100 (.011)*	.007 (.026)	-.148 (.019)*
TECH17	.046 (.017)	.015 (.030)	.064 (.068)	-.022 (.018)	n.a.	-.105 (.022)*
TECH18	.000 (.015)	.053 (.017)*	-.194 (.046)*	.208(.023) *	-.418 (.064)*	n.a.
Relative wage:						
$\frac{WLSM}{WMAN}$	.179 (.035)*	-.723 (.167)*	.059 (.032)	.114 (.029)*	.143 (.112)	.262 (.060)*
$\frac{WAUT}{WMAN}$	.217 (.032)*	.227 (.081)*	-.622 (.072)*	-.360 (.040)*	-.696 (.122)*	-.515 (.084)*
DQ2	.051 (.007)*	.119 (.013)*	.012 (.012)	-.034 (.008)*	.016 (.021)	-.053 (.016)
DQ3	-.044 (.007)*	.006 (.012)	-.002 (.011)	-.012 (.008)	-.012 (.021)	-.045 (.016)
DQ4	-.054 (.008)*	.022 (.020)	-.003 (.013)	.008 (.009)	-.057 (.026)*	-.025 (.023)
TREND	-.008 (.001)*	-.047 (.003)*	-.020 (.002)*	.024 (.001)*	-.021 (.003)*	.054 (.003)*
DC	-.045 (.012)*	-.271 (.090)*	.076 (.022)*	.034 (.014)*	-.043 (.037)	-.100 (.034)
$\hat{\sigma}$	.175	.188	.334	.218	.555	.424
R^2	.969	.967	.904	.964	.852	.876
Sample size	6849	2952	6654	6848	6039	6079

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

Table 10

Output Coefficients and Variabilities from a Quadratic Model

(robust standard errors in parentheses)

	Coefficient on q	Coefficient on q^2	Mean output variability	(25^{th} , 75^{th}) Percentile Sorted by output level
FD - IV Estimator				
Flat Sorting				
Manual (05)	.960 (.045)*	.002 (.019)	.971	(.969, .973)
FSM- All (11)	.916 (.042) *	.009 (.021)	.967	(.956, .976)
FSM 881 (19)	.948 (.039)*	.008 (.014)	.995	(.986, 1.004)
FSM 1000 (20)	.274 (.167)	.550 (.449)	3.682	(3.17, 4.17)
Letter Sorting				
Manual (06)	.996 (.037)*	.011 (.011)	1.095	(1.083, 1.108)
LSM (02)	.983 (.118)*	.463 (.364)	5.300	(4.700, 5.991)
OCR (01)	.974 (.055)*	.005 (.017)	1.022	(1.016, 1.028)
BCS- All (10)	.981 (.043)*	.012 (.015)	1.096	(1.082, 1.112)
BCS (17)	.850 (.095)*	.024 (.041)	1.082	(1.055, 1.111)
DBCS (18)	.998 (.077)*	.025 (.089)	1.233	(1.206, 1.263)
FE - IV Estimator				
Flat Sorting				
Manual (05)	.928 (.132)*	-.013 (.034)	.862	(.876, .846)
FSM- All (11)	1.095 (.138)*	-.033 (.024)	.908	(.944, .873)
FSM 881 (19)	.802 (.105)*	.000 (.021)	.803	(.803, .803)
FSM 1000 (20)	.847 (.647)	-.016 (.092)	.744	(.760, .730)
Letter Sorting				
Manual (06)	.436 (.225)	.070 (.027)*	1.094	(1.014, 1.185)
LSM (02)	2.265 (.678)*	-.136 (.075)	.997	(1.171, .795)
OCR (01)	1.094 (.318)*	-.026 (.038)	.847	(.876, .814)
BCS-All (10)	1.546 (.783)*	-.041 (.033)	1.165	(1.211, 1.112)
BCS (17)	-.208 (.944)	.109 (.105)	.841	(.722, .975)
DBCS (18)	1.881 (.817)*	-.072 (.085)	1.194	(1.274, 1.105)

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

Table 11

Delivery Points as a Labor Demand Shifter

	Coefficient on *DPT*		Output Variability	
	FD/IV	FE/IV	FD/IV	FE/IV
Flat Sorting				
Manual (05)	.007 (.004)	-7.180 (3.39)	.963 (.045)	.908 (.073)
FSM-All (11)	.002 (.004)	.717 (1.41)	.923 (.041)	.957 (.061)
FSM 881 (19)	-.006 (.007)	-7.708 (2.00)*	.950 (.040)	.836 (.052)**
FSM 1000 (20)	.005 (.011)	-11.748 (5.68)	.383 (.144)**	.796 (.250)
Letter Sorting				
Manual (06)	-.000 (.002)	-4.132 (1.40)*	.985 (.036)	1.009 (.051)
LSM (02)	.012 (.004)*	5.353 (2.96)	.994 (.153)	1.151 (.190)
OCR (01)	-.013 (.005)*	-11.002 (4.49)*	.972 (.055)	.903 (.087)
BCS- All (10)	.010 (.003)*	5.941 (1.81)*	.964 (.041)	1.191 (.056)**
BCS (17)	-.003 (.010)	2.454 (4.51)	.834 (.097)	.654 (.163)**
DBCS (18)	.004 (.005)	18.154 (4.77) *	.971 (.073)	1.156 (.165)

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

Table 12

Output Variabilities with Lagged Adjustment

(Robust standard errors in parentheses)

FD/IV estimator	q_t - q_{t-1}	q_{t-1} - q_{t-2}	q_{t-2} - q_{t-3}	q_{t-3} - q_{t-4}	Sum
Flat Sorting					
Manual (05)	.966 (.054)*	.098 (.070)	.007 (.061)	.026 (.033)	1.097 (.140)
FSM-All (11)	.915 (.041)*	-.020 (.042)	.035 (.035)	.007 (.030)	.937 (.088)
FSM 881 (19)	.934 (.039)*	-.035 (.040)	-.003 (.039)	-.011 (.042)	.885 (.109)
FSM 1000 (20)	.405 (.170)	.106 (.109)	.115 (.077)	.038 (.059)	.664 (.295)
Letter Sorting					
Manual (06)	.983 (.037)*	.018 (.040)	.045 (.038)	.044 (.034)	1.091 (.108)
LSM (02)	.889 (.076)*	.138 (.077)	.192 (.067)*	.178 (.056)*	1.396 (.203)
OCR (01)	.949 (.062)*	-.049 (.067)	-.035 (.061)	.082 (.049)	.948 (.165)
BCS -All (10)	1.036 (.049)*	.176 (.052)*	.097 (.043)*	.087 (.036)*	1.396 (.117)
BCS (17)	.828 (.101)*	-.060 (.103)	-.105 (.098)	.004 (.077)	.667 (.295)
DBCS (18)	1.179 (.101)*	.505 (.117)*	.142 (.083)	.075 (.048)	1.901 (.199)

FE/IV estimator	q_t	q_{t-1}	q_{t-2}	q_{t-3}	Sum
Flat Sorting					
Manual (05)	.802 (.108)*	.036 (.095)	.044 (.081)	.072 (.075)	.954 (.104)
FSM-All (11)	.957 (.066)*	-.022 (.061)	.026 (.051)	.017 (.039)	.977 (.071)
FSM 881 (19)	.821 (.060)*	-.104 (.052)	.032 (.060)	.036 (.046)	.784 (.063)
FSM 1000 (20)	.710 (.245)*	.282 (.134)	.148 (.130)	.094 (.090)	1.234 (.336)
Letter Sorting					
Manual (06)	.931 (.050)*	-.035 (.046)	.086 (.049)	.092 (.039)	1.074 (.060)
LSM (02)	.807 (.117)*	.154 (.101)	.244 (.085)*	.249 (.086)*	1.454 (.186)
OCR (01)	.879 (.089)*	-.014 (.089)	-.053 (.081)	.070 (.075)	.882 (.109)
BCS-All (10)	1.101 (.064)*	.158 (.056)*	.041 (.053)	.024 (.051)	1.325 (.072)
BCS (17)	.493 (.220)*	.121 (.204)	.101 (.141)	.220 (.154)	.935 (.230)
DBCS (18)	1.160 (.215)*	.250 (.205)	.035 (.320)	-.173 (.256)	1.273 (.187)

* reject that the coefficient is equal to zero at the .01 level with a two-tailed test

Table 13

Output Variabilities Using TPF by Sorting Operation as the Output Measure

(robust standard errors in parentheses)

				Instrumental Variables	
	OLS	Fixed Effects	First Difference	Fixed Effects	First Difference
Flat Sorting Operations					
Manual (05)	.815* (.012)	.638* (.007)	.454* (.036)	.894 (.061)	.913 (.037)
FSM- All (11)	.969 (.016)	.676* (.010)	.537* (.089)	.920* (.027)	1.021 (.032)
FSM881 (19)	.951* (.014)	.641* (.010)	.487* (.082)	.944 (.037)	1.014 (.034)
FSM1000 (20)	.838* (.020)	.660* (.036)	.672* (.043)	1.939 (.910)	.997 (.275)
Letter Sorting Operations					
Manual (06)	.794* (.009)	.442* (.008)	.475* (.040)	.901* (.046)	.970 (.046)
LSM (02)	.934* (.008)	.871* (.009)	.842* (.032)	1.060 (.119)	1.091 (.121)
OCR (01)	.777* (.019)	.438* (.016)	.449* (.067)	.988 (.092)	1.028 (.058)
BCS -All (10)	.937 (.053)	.640* (.014)	.324* (.052)	1.197* (.055)	1.060 (.050)
BCS (17)	.957* (.021)	.722* (.012)	.527* (.055)	.759 (.138)	1.090 (.097)
DBCS (18)	1.019 (.036)	.718* (.018)	.269* (.055)	1.613* (.225)	1.118 (.101)
Priority-Parcel Sorting Operations					
Manual Parcel (07)	.521* (.025)	.216* (.022)	.122* (.032)	.305* (.325)	.459 (.253)
SPBS (12)	.676* (.053)	.390* (.013)	.482* (.087)	.970 (.131)	.642 (.143)
Manual Priority (08)	.719* (.017)	.664* (.013)	.454* (.035)	1.137 (.138)	.813 (.117)

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

Table 14

Output Variabilities For Priority Mail Sorting

(robust standard errors in parentheses)

(Sample consists of 158 plants that sort priority mail in every postal quarter from1994-2000)

				Instrumental Variables	
	OLS	Fixed Effects	First Difference	Fixed Effects	First Difference
Priority Mail Sorting Operations					
Manual (08)	.806* (.022)	.844* (.034)	.424* (.070)	1.105 (.109)	.893 (.076)
SPBS (04)	.722* (.060)	.513* (.077)	.313* (.093)	1.898 (.422)	1.054 (.084)

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