Postal Rate Commission
Submitted 11/20/2006 4:21 pm
Filing ID: 55071
Accepted 11/20/2006

USPS-RT-5

BEFORE THE
POSTAL RATE COMMISSION
WASHINGTON, D. C. 20268-0001

POSTAL RATE AND FEE CHANGES, 2006	Docket No. R2006–1

REBUTTAL TESTIMONY
OF
A. THOMAS BOZZO
ON BEHALF OF THE
UNITED STATES POSTAL SERVICE

Table of Contents

List of Tables

Category 2 Library Reference to be Sponsored with USPS-RT-5

USPS-LR-L-192 Materials Relating to the Rebuttal Testimony of A. Thomas Bozzo (USPS-RT-5)

Autobiographical Sketch

My name is A. Thomas Bozzo. I am a Vice President with Laurits R. Christensen Associates (LRCA), which is an economic research and consulting firm located in Madison, Wisconsin. My education and experience are described in detail in my direct testimony, USPS-T-12 and USPS-T-46.

Purpose and Scope of Testimony

The purpose of this testimony is to rebut criticisms of the Postal Service's econometric estimates of volume-variability factors for mail processing labor, and of the underlying economic theory and econometric methods, found in the testimonies of witnesses Haldi (VP-T-2), Neels (UPS-T-1) and Roberts (OCA-T-1).

Associated with my testimony is the Category 2 Library Reference USPS-LR-L-192, which contains the background material for the analyses reported in this testimony. The accompanying CD-ROM contains electronic versions of the econometric estimation code and data used for the analyses presented herein.

I. Dr. Haldi Mischaracterizes Economic Cost Theory Pertaining to Economies of "Scale" and the Treatment of Non-Volume-Variable Costs.

I.A. Dr. Haldi's Discussion of Economies of "Scale" and "Density" Mischaracterizes the Relevant Economic Theory and the Postal Service Models and Should Be Disregarded.

I.A.1. Dr. Haldi Admits that Cost Elasticities are an Appropriate Aim of the Analysis.

In large part, Dr. Haldi's criticism of the Postal Service models as failing to provide information on "economies of scale" is beside the point of the analysis. The primary purpose of the analysis is to measure cost elasticities required to implement the mail processing volume-variable cost calculations for the sorting cost pools that are the subject of the analysis. USPS-T-12 at 33. Dr. Haldi agrees that cost elasticities are, indeed, inputs to volume-variable costs. Tr. 23/8624. I do not, for that matter, actually claim that the elasticities I measure provide information on economies of "scale," though as discussed below, they do bear on the related concept of economies of "density."

I.A.2. Dr. Haldi's Claim that Cost Elasticities Provide Little Information on the Presence of "Scale Economies" Is Incorrect.

Central to Dr. Haldi's discussion of "scale" economies is the assertion that the Postal Service mail processing variability models specifically are unable to provide much information on the presence or absence of "scale" economies in sorting operations. VP-T-2 at 17-31. Dr. Haldi's claim is inconsistent with the relevant economic theory, which provides that considerable information regarding

scale economies may be obtained from cost functions, or related factor demand functions.

When there exist economies of scale, an expansion of the firm's inputs lead to a proportionally larger increase in output. Or, equivalently, a given increase in output can be achieved with a less than proportionate increase in inputs. The connection to cost elasticities follows directly from the fact that if inputs increase less than proportionately with a given output increase, costs (i.e., of the inputs) will also increase less than proportionately. Robert G. Chambers, Applied Production Analysis, Cambridge University Press 1986, p. 71-72.

Dr. Haldi attempts to suggest that the connection between cost elasticities (or economies of "size") and economies of scale is limited to certain special conditions. Tr. 23/8625-8626. Dr. Haldi misreads Prof. Chambers's monograph in reaching this conclusion. The conditions Dr. Haldi mentions (cost minimization, "hometheticity" of the cost function[1]) are conditions under which scale and size economies coincide *exactly*. Response to USPS/VP-T2-13, Tr. 8601-8602. Prof. Chambers notes that, as a rule, economies of size are of greater interest because firms' cost responses to changes in output are more important than considering the response of output to input combinations that need not result from the firm's decision processes. Robert G. Chambers, op. cit., at 74.

[1] It should be noted that many functional forms commonly used in applied cost analysis, including the Cobb-Douglas, quadratic, and translog, either are homoethetic by construction or can be restricted to be homothetic; it is a common and useful assumption for applied analysis.

As I noted above, the labor demand analysis bears on economies of "density" rather than economies of scale. In analysis of network industries, economies of "density" distinguish the cost effects of increasing outputs from those of increasing both outputs and the network served by the firm. The concept is germane to analysis of Postal Service costs because it must serve delivery points that are growing steadily with population, while mail volumes have been relatively flat.

Dr. Haldi seems to consider it surprising that there could be economies of "density" for large and small facilities alike (response to USPS/VP-T2-15, Tr. 23/8605-8606), but this is in fact no surprise. It would be expected that facilities of all sizes would have operations that are not operating exactly at capacity, so that marginal increases in volume would spread non-volume-variable costs (or, more generally, less-than-fully volume-variable costs) over a larger output base. This point has long been emphasized by the Postal Service's operations testimony. It is, in fact, not unlike observing that an airline can lower its average costs by filling otherwise-empty seats on both a 150-seat jet and on a 400-seat jumbo. The potential existence of economies of density is not, in this respect, dependent on the scale of operations.

I.A.3. Dr. Haldi's Claim that the Postal Service Models Omit Facility Size Variables is False

An especially curious claim Dr. Haldi makes is:

> [U]ntil witness Bozzo either disaggregates and analyzes his data according to plant size, ***or introduces explicit variables for plant size***, inferences on cost elasticities developed by witness Bozzo

> are not likely to provide insight on quantities such as scale elasticities. Response to USPS/VP-T2-14(b), Tr. 23/8604; emphasis added.

Technically, Dr. Haldi is incorrect to suggest that it is necessary to incorporate plant size measures or to conduct a disaggregated analysis to measure the elasticity of size, or to obtain insight into "scale" economies from it. However, this is secondary to the issue that Dr. Haldi has criticized models with which he is evidently not familiar.

Asked what he would consider to be variables representing plant size, Dr. Haldi agreed that measures of capital stock, capital input, and/or the delivery network could serve as plant size variables. Tr. 23/8629-8630. Dr. Haldi, therefore, seems to have failed to notice that the Postal Service labor demand models actually include *both* capital and delivery point variables as plant size measures. USPS-T-12 at 52, 53 (equation 16), and 54 (equation 17). For that matter, Dr. Neels's and Prof. Roberts's models also include what Dr. Haldi considers facility size measures of some sort. Response to USPS/UPS-T1-5, Tr. 23/8467; OCA-T-1 at 19. There could exist labor demand models subject to Dr. Haldi's critique, but my models and those of the intervenor witnesses are not among them.

I.B. Dr. Haldi's Theories of Setup Costs Have Negligible Bearing on Costs for CRA Subclass Categories.

Dr. Haldi describes at some length theories according to which he claims that some non-volume-variable setup and take-down costs in sorting operations may constitute incremental costs for certain classes of mail and other product

categories. VP-T-2 at 39-47. Dr. Haldi suggests that, after accounting for non-volume-variable costs that are attributable as incremental costs, the Commission could in principle find itself with similar results to the present distribution of costs under the Commission's 100 percent volume-variability assumption. *Id.* at 55.

Dr. Haldi's argument fails for several interrelated reasons. First, the setup costs clearly are not volume-variable (i.e., "attributable" as marginal costs). Second, a significant portion of the setup costs Dr. Haldi discusses cannot be assigned to any class of mail as incremental cost, and even those costs that may be class-specific are not, in general, the incremental costs of any subclass. Thus, methods intended to distribute pools of volume-variable cost to subclasses in order to represent marginal costs will have little application to these non-volume-variable costs. Finally, it should be noted that the Postal Service incremental cost model includes "inframarginal" variable costs, so the possibility that setup costs may represent costs that may vary with volume in some respects, but not "on the margin," is already incorporated in the incremental cost estimates.

I.B.1. Setup Costs are Not Marginal Costs, and a Significant Portion of Setup Costs is "Fixed," Particularly in Incoming Schemes.

Witness McCrery noted that sort schemes tend not to vary with volumes on the margin:

> [E]ven if volume declines, perhaps due to a rate increase, all of the sort schemes must still be run. This results in what [witness Kingsley, Docket No. R2000-1, USPS-T-10] calls the "schemes effect". This effect is, in my judgment, a major reason why

> workhours commonly vary less than volume… in individual mail processing operations. USPS-T-42 at 35, lines 22-26.

Witness McCrery further notes that:

> Out of the hundreds of schemes run each day, a few very large schemes (e.g., the initial outgoing and incoming schemes, i.e., the "primaries") may be run on multiple sorters due to time constraints. But for the vast majority of schemes, this is neither necessary nor desirable. *Id.* at 36, lines 11-14.

Witness McCrery's account explains why scheme change time should not be considered volume-variable to any appreciable degree. Small changes in volume on the margin are insufficient to require the addition or elimination of scheme runs. Dr. Haldi agrees that large incremental changes in volume would be necessary to add or eliminate schemes. Response to USPS/VP-T2-19, Tr. 23/8613. Dr. Haldi also correctly counsels that care should be taken not to confuse volume-variable costs with incremental costs such as inframarginal variable costs. Response to USPS/VP-T2-19, Tr. 23/8612-8615. Thus, scheme changes are not appropriately considered to be marginal (i.e., volume-variable) costs.

Witness McCrery shows, further, that many schemes, notably incoming secondary schemes (*e.g.*, delivery point sequencing) are very rarely run on multiple sorters. Response to MPA-ANM/USPS-T42-22(d)-(e), Tr. 11/2895-2897. In such cases, the setup costs are "fixed" with respect to "inframarginal" volume changes as well as small changes on the margin. Even in cases where schemes are run on multiple machines, avoiding the full setup cost would require eliminating all volumes from the scheme, so a portion of those costs are appropriately considered "fixed" as well.

As I explain below, the Postal Service incremental cost model appropriately incorporates inframarginal variable costs in the incremental cost estimates, and the "fixed" costs are generally not part of any subclass incremental costs.

I.B.2. The Postal Service's Incremental Cost Model Appropriately Accounts for Inframarginal Variable Costs.

Certainly, a portion of setup costs and other non-volume-variable mail processing costs may represent what may be called "inframarginal" variable costs—that is, costs which may vary given some incremental change in volume that is more than unit change (marginal cost, literally) but less than the entire product volume (as in "intrinsic" costs). Dr. Haldi counsels care in distinguishing marginal, inframarginal, and "intrinsic" costs. Response to USPS/VP-T2-19, Tr. 23/8614-8615. In fact, witness Pifer's incremental cost model incorporates inframarginal costs via a constant-elasticity approximation to component cost functions. USPS-T-18 at 8; see also Docket No. R2000-1, USPS-T-22. Witness Pifer's treatment is conceptually correct and clearly superior to forcing activities with non-volume-variable (but also not totally "fixed") costs into volume-variable cost pools.

I.B.3. In "Class-Specific" Schemes, "Fixed" Setup Costs Are Not Attributable To Subclasses as Incremental Costs; Mixed-Class Schemes are Not Attributable at the Class or Subclass Level.

Dr. Haldi makes some effort to distinguish schemes that may be class-specific with schemes in which classes of mail are routinely merged. See, e.g.,

VP-T-2 at 43-47. However, in neither case are the "fixed" costs appropriately considered "intrinsic" costs at the subclass level, and in the case of schemes where classes of mail are merged, there is no basis for attribution as incremental costs even at the class level.

If the setup costs are "fixed" (*i.e.*, avoidable only if the entire scheme is eliminated), then it is necessary to withdraw all volume from the scheme in order for the setup costs to be avoided. But since even "class-specific" schemes generally merge significant volumes of mail from multiple subclasses, eliminating the total volume of mail from any subclass will not be sufficient to eliminate the scheme. Where the setup cost is not avoidable with the (hypothetical) withdrawal of the subclass, it cannot be considered to be part of the subclass incremental cost. Response to USPS/VP-T2-23b, Tr. 23/8622.

When multiple classes are merged in a scheme, the setup costs are not incremental costs at the class level, either. In such cases, even the elimination of the volume of an entire class of mail will be insufficient to lead to avoidance of the setup costs. Response to USPS/VP-T2-22, Tr. 23/8620. The costs are "institutional costs" without a causal basis for assignment to class or subclass.

I.B.4. Dr. Haldi's Concerns are Moot in Key Outgoing Operations Using Postal Service Variabilities.

Dr. Haldi agrees that his concerns are mooted by unit variabilities – in such cases, the affected operations have no non-volume-variable costs to deal with. Response to USPS/VP-T2-21; Tr. 23/8619. While Dr. Haldi means to suggest that the Commission's assumptions make the incremental cost

assignment problems go away, his comment conceivably applies to any variability method. In fact, in many of the operations where Dr. Haldi stakes his case, the Postal Service variabilities are statistically 100 percent, notably outgoing BCS and AFSM 100 operations.

Conversely, significant non-volume-variable costs appear in automated incoming BCS sorting, where the evidence suggests that the setup costs are predominantly "fixed" (see witness McCrery's response to MPA-ANM/USPS-T42-22, op. cit.) and the routine mixing of classes implies that the setup costs are not, in fact, incremental to any subclass or class.

I.B.5. Non-Volume-Variable Costs in Sorting Operations are Insufficient to Change the Outcome of Class-Level Incremental Cost Tests.

Dr. Haldi fails to show, and indeed cannot show, that even assigning the entirety of non-volume-variable costs in sorting operations to a particular class of mail – let alone the subset of setup costs that might actually be caused by a class of mail as intrinsic costs – would affect the outcome of the affected class-level incremental cost tests. In the Postal Service method, the total non-volume-variable cost in sorting operations totals $721 million. In contrast, the difference between First-Class Mail revenue and incremental cost is $18,891 million and that between Standard Mail revenue and incremental cost is $10,240 million, according to Dr. O'Hara. (USPS-T-31 at 19, 28, 30.)

Adding the entire non-volume-variable cost to either First-Class Mail or Standard Mail, which as discussed above would be grossly inappropriate, the incremental cost tests would still show no cross-subsidy. As a practical matter,

the magnitudes of non-volume-variable costs in sorting operations are insufficient to lead to problems even if Dr. Haldi's arguments are given maximum credence. At best, Dr. Haldi describes a theoretical problem that careful analysis shows not to have any practical significance for the ratemaking scheme's cost tests.

II. Major Errors Lead Dr. Neels's Analysis of IOCS Activity Data for MODS Cost Pools to Greatly Overstate The Possibility of Clocking Errors in Sorting Operations.

Dr. Neels presents partial results from a crosswalk of clerk and mail handler activities as recorded in IOCS with MODS cost pools, which he purports to show a "serious problem" of "workers clocked into the wrong sorting operation." UPS-T-1 at 15, esp. Table 3. However, Dr. Neels's IOCS analysis is faulty in both the calculations and interpretation of the results. The shortcomings of Dr. Neels's analysis includes failure to include several significant IOCS activities needed to ensure consistency between the IOCS activities and MODS cost pools, and inappropriately interpreting the presence of other activities that are legitimate components of the cost pools as indicative of misclocking. I show below that a correct reading of the IOCS data shows the IOCS and MODS operations to be substantially in accord with each other.

II.A. Dr. Neels Admits to Mistakenly Omitting Several Important Operations from His Analysis.

Dr. Neels intended his analysis to show inconsistency of IOCS activities and MODS cost pools for the operation groups used in the econometric variability analysis. Tr. 23/8543. However, his analysis failed to record several major activities as being correctly clocked: MPBCS activities in the D/BCS cost pool (which encompasses MPBCS and DBCS operations), LIPS activities in the SPBS cost pool (which encompasses SPBS and LIPS operations), and "allied labor" for collection mail, such as dumping containers and culling mail (a component of the

Cancellation cost pool). Response to USPS/UPS-T1-6(a)-(f); Tr. 23/8470-8471. Tallies for these IOCS activities appeared in Dr. Neels's Table 3 as representing "Non-Sorting Activities," which the Table 3 title implies (incorrectly) represent clocking problems, even though the activities are actually consistent with the clocked sorting operation. Dr. Neels admits that the correct treatment of these tallies would regard them as being consistent with the clocked operation. Response to USPS/UPS-T1-6(g)-(h); Tr. 23/8470-8472.

II.B. Dr. Neels's Analysis Inappropriately Implies Certain "Overheads" and Incidental Other Activities Constitute Clocking Errors.

In addition to the IOCS activities Dr. Neels failed to correctly treat as part of the same sorting operation as the MODS cost pool, Dr. Neels's Table 3 also inappropriately implies a variety of other IOCS "non-sorting" activities are emblematic of clocking problems. This is not a correct interpretation of the data.

As I explained in my direct testimony, the MODS cost pools include a certain amount of labor hours in "overhead" activities (mainly breaks and clocking in or out), "quasi-allied labor," and incidental "administrative" activities. USPS-T-12 at 26-27.

II.C. A Corrected Analysis Shows MODS and IOCS Sorting Activities to be Largely Consistent.

To show the actual degree of consistency between MODS and IOCS, I corrected Dr. Neels's major errors, discussed above. I also correctly classified certain other tallies indicating a sorting or cancellation activity as relating to the

appropriate sorting activity in some cases where Dr. Neels's code allowed those tallies to revert to the non-sorting activity categories.

The corrected data for sorting activities are shown in Table 1, below. The IOCS sorting activity is fully consistent (*i.e.*, same activity) with the MODS cost pool for 93 percent of the sorting activity tallies. It should be noted that Dr. Neels's presentation of the data, apart from the errors requiring correction, does not show that the largest cost pools have relatively high fractions of tallies where the sorting activity is consistent with the MODS cost pool.

Table 1. Comparison of MODS and IOCS Activity Responses

MODS Cost Pool	IOCS Activity: Same Activity	Related Sorting Activity	Other Sorting Activity	% Same or Related Activity
D/BCS	1,335,160	27,400 /1	16,298	99%
OCR	141,877	43,227 /1	3,398	98%
AFSM/100	483,935	19,924 /2	11,382	98%
FSM/1000	190,042	13,290 /2	7,732	96%
SPBS	512,581	13,072 /3	33,715	94%
Manual Flats	212,092	13,364 /4	4,232	98%
Manual Letters	828,276	30,774 /4	13,559	98%
Manual Parcels	55,934	16,321 /4	3,789	95%
Manual Priority	187,794	30,008 /4	11,068	95%
Cancellations	247,574	-	11,209	96%
Total	4,195,265	207,380	116,384	97%

Notes:
1/ Other automated letter sorting
2/ Other automated flat sorting
3/ Other mechanized package sorting
4/ Other manual piece sorting

It was clear from the data underlying Dr. Neels's analysis that tallies for "Other sorting activities" were concentrated in related activities – *e.g.*, automated

equipment handling the same shape of mail.[2] Taking this into account further narrows the scope of potential misclocking problems – there is no evidence from IOCS that it is necessary to consider the presence of significant hours misclocked to operations unrelated to the MODS cost pool. It should be noted that IOCS, while generally reliable, is not a totally error-free data system, and any inaccuracies would contribute to the anomalies. Regardless of the cause, Dr. Neels's analysis clearly exaggerates the extent of clocking anomalies.

II.D. IOCS "Administrative" Tallies in MODS Mail Processing are Not Problematic.

As an entrée to his discussion of purported MODS clocking errors, Dr. Neels cites past controversies regarding the appropriate treatment of MODS mail processing tallies with IOCS "administrative" operation codes. UPS-T-1 at 14. The Commission has been treating the costs associated with certain IOCS tallies as representing general administration rather than as mail processing-related administration, based in part on the testimony of UPS witnesses that the "traditional" assignment of clerk and mailhandler costs to the Cost Segment 3 components should be retained.

While Dr. Neels suggests that a similar treatment might still be warranted in this proceeding, he not only fails to provide any evidence that the traditional treatment is warranted for administrative tallies in "Function 1" operations, but in fact he admits that various administrative activities are defined such that it is not

[2] In some cases, IOCS data collectors may indicate a consistent category in the remarks rather than in the selected option; I considered such tallies "related activities" for the purposes of this analysis.

at all surprising to encounter them in mail processing operations. Response to USPS/UPS-T1-28, Tr. 23/8510; see also Tr. 23/8548-8551. In fact, as will be shown below, the great majority of "administrative" tallies fall within a subset of MODS operations expressly defined for administrative and miscellaneous activities. The remainder occur in mail processing operations in a manner consistent with infrequent, incidental incurrence. This should satisfy the Commission's indication, expressed in Docket No. R2000-1, that the Postal Service should provide evidence to link the tallies to mail processing operations. PRC Op., Docket No. R2000-1, at ¶3014.

II.D.1. The Small Fractions of "Administrative" Tallies in Sorting Operations are Not Unexpected.

In the MODS sorting operations covered by the econometric variability analysis, tallies with "administrative" operation codes constitute a small fraction of costs. Most of those are associated with clocking in or out activities, which IOCS processing assigns activity code 6522 and an administrative operation code regardless of the clocked-in MODS operation.[3] Dr. Neels agrees that clocking in and out is a component of the sorting operations. Response to USPS/UPS-T1-6, Tr. 23/8470-8472; Response to USPS/UPS-T1-27, TR. 23/8508-8509. The remaining costs associated with administrative tallies in sorting operations are 0.7 percent of the total pool costs. The correct disposition notwithstanding, Dr.

[3] The traditional treatment of clocking in and out costs in Cost Segment 3.1 assigned portions of the costs back to the mail processing and window service components to correctly reflect the causation of those costs.

Neels agrees that these costs are insignificant. Response to USPS/UPS-T1-27, Tr. 23/8508-8509; Tr. 23/8552.

It should be recognized that the nature of many "administrative" activities is such that they actually would be incurred, in small amounts, in mail processing operations including sorting operations. Activities included in the IOCS "administrative" category include "Data Collection & Processing Activities… Training, Quality Control/Revenue Protection… [and] Union Business" (text of IOCS Question 18B, "Operational Area," USPS-LR-L-9), any of which may be observed in small amounts for correctly clocked mail processing employees. Therefore, these should not be viewed as emblematic of clocking problems.[4]

II.D.2. The Vast Majority of "Administrative" Tallies are in MODS Operations Expressly Defined for Mail Processing Administration.

In the MODS system, a number of operations are defined expressly for administrative and miscellaneous activities. It should not be at all surprising that the bulk of the mail processing tallies with "administrative" operation codes appear in those operations. Dr. Neels agreed that administrative tallies would be expected in those operations. Response to USPS/UPS-T1-28, Tr. 23/8510. However, he evidently was not aware of the extent to which the tallies were concentrated in those operations, as he opined that the fraction of tallies in those

[4] Dr. Neels raises the issue of why the employees do not re-clock into other operations in his response to USPS/UPS-T1-4(d) (Tr. 23/8465). He answers his own question subsequently by observing the fairly substantial clocking in and out costs incurred by the Postal Service, response to USPS/UPS-T1-27, (Tr. 23/8509); having employees re-clock for very small amounts of incidental "administrative" activities would be costly and inefficient.

operations was lower than 73 percent (Tr. 23/8551), which is the actual fraction, as shown in Table 2, below. This should not be taken as a ceiling on the fraction of correctly assigned administrative tallies. The list of operations in Table 2 was selected for clarity of the operation names, and excludes operations that would have significant "administrative" components, for example mail acceptance operations and dock expediter work, both of which involve administrative tasks such as record-keeping.

Table 2. IOCS "Administrative" Tallies for Select Operations, BY 2005.

Op Code	Operation Description	Admin tallies ($000), excluding a/c 6522
340	STANDBY - MAIL PROCESSING	4,185
341	QWL COORDINATOR - NONSUPER EMPS	2,262
547	SCHEME EXAMINERS	8,127
554	OFFICE WORK & RECORDS-MAIL PROC	71,454
555	OFFICE WORK & RECORDS-MAIL PROC	19,001
560	MISC ACTIVITY-MAIL PROC	30,755
561	MISC ACTIVITY-MAIL PROC	7,600
562	MISC ACTIVITY-MAIL PROC	5,722
563	MISC ACTIVITY-MAIL PROC	9,191
564	MISC ACTIVITY-MAIL PROC	7,296
577	PREP & VERIFY DELV BILLS-INTERNAT	410
607	STEWARDS - CLERKS - MAIL PROC	46,304
612	STEWARDS-MAIL HANDLER-MAIL PROC	19,778
630	MEETING TIME-MAIL PROC	5,177
677	ADMIN & CLER - PROCESSING & DISTRIB	30,637
681	ADMIN & CLER - PROC & DIST INTERNTL	1,986
697	ADM & CLER-MAIL.REQ & BUS.MAIL ENT	4,998
	Subtotal Above Operations	274,883
	Total MODS Mail Processing "Administrative" Tallies	377,113
	Percent in listed MODS operations	73%

II.D.3. The "Administrative" Tallies are Appropriately Included in Cost Segment 3.1.

Given that a large majority of IOCS "administrative" tallies in MODS mail processing operations specifically associated with mail processing, and indeed all are incurred at mail processing facilities, the Commission should determine that there is sufficient evidence to consider these costs as part of the mail processing component, consistent with the Postal Service treatment, rather than as general administrative costs in Cost Segment 3.3.

III. Dr. Neels Greatly Overstates the Extent of MODS Data Problems.

Dr. Neels raises an array of claimed MODS data problems as obstacles to implementing a MODS-based volume-variability analysis (UPS-T-1 at 15-20), even as Prof. Roberts has instead pursued robust econometric methods that permit allow him to recommend some MODS-based econometric variability models. OCA-T-1 at 52. In this proceeding, Dr. Neels has pursued analyses which seem to suggest that the Postal Service data for automated operations is no better than, if not actually worse than, its data for manual operations. UPS-T-1 at 22-23. This result is not only counterintuitive, considering that MODS workload measurement processes for automated operations employ machine counts that are reliable in principle, but as witness Oronzio points out, directly counter to the experience of Postal Service field managers and operations experts. USPS-RT-15 at 10-12.

Dr. Neels's effort to cast every conceivable MODS data anomaly as a fatal error contradicts his previous testimony, in particular that from Docket No. R97-1, where Dr. Neels criticized Prof. Bradley for eliminating too many "usable" observations from his samples. Docket No. R97-1, UPS-T-1 at 46; Tr. 28/15632. Dr. Neels also depends on a variety of misinterpretations and misuses of the MODS data. In particular, Dr. Neels's analysis conflates conceptual definitions of MODS measures with the actual measurement methods, conducts screens that defeat the purpose of pooling MODS operations, and ignores whether the claimed anomalies are relatively small or relatively large. As such, Dr. Neels's

own present handling of anomalies falls squarely in the sights of his earlier critiques of the Postal Service. The cumulative effect is to greatly overstate the econometrically relevant MODS data issues.

III.A. Dr. Neels's Current Testimony is Inconsistent With His Past Approach to MODS Data "Anomalies."

Dr. Neels's Docket No. R97-1 criticism of Prof. Bradley's models, in criticizing both what he considered to be excessive data screening and a failure to deal with measurement error in MODS piece handlings (Docket No. R97-1, UPS-T-1 at 22; Tr. 26/15608; p.26-27, Tr. 27/15612-15613), did at least point towards suggest a reasonable general approach to handling the MODS data.[5] Specifically, Dr. Neels's earlier testimony could have been read as appropriately suggesting that researchers should tread lightly on data that are merely "anomalous," and should use appropriate econometric techniques to deal with unavoidable measurement errors. This is basically the method Prof. Roberts and I have adopted in our respective studies. Prof. Roberts notes that it is important to eliminate gross data errors where they may be identified (response to USPS/OCA-T1-38, Tr. 23/8372), but also notes the considerable utility of the instrumental variables technique for dealing with issues such as measurement error in MODS FHP. Roberts 2002 at 55.

Dr. Neels's current analysis of the MODS data suggests he has discarded his original data screening philosophy. The proof is in Dr. Neels's failure to

[5] Many of the specific details of Dr. Neels's critique were wanting; see Docket No. R2000-1, USPS-RT-7 at 14-26 (Tr. 46E/22049-22061).

discern the magnitudes of certain potential errors (response to USPS/UPS-T1-11, Tr. 23/8484), or to determine whether others even represented errors at the level of analysis employed in Dr. Neels's or the Postal Service models. Response to USPS/UPS-T1-9, Tr. 23/8479. Indeed, Dr. Neels's goal seems to be to label as many erroneous observations he can, and to throw up his hands at the mysteries of the MODS data. Tr. 23/8568-8572. This contrasts with Prof. Roberts's approach, which eliminates some erroneous data but in which data issues ultimately are not insurmountable obstacles to recommending MODS-based models.

III.B. Dr. Neels's Critiques Fail to Correctly Account for MODS Measurement Methods.

While it is common to discuss MODS as a monolithic system, the MODS system uses separate measurement systems for its key data elements: workhours are derived from clock rings, automated TPF and TPH are derived from piece counts on the sorting equipment, FHP are weighed and converted to pieces, and manual TPH are projected from FHP and mail downflows. It is possible for data "anomalies" to arise solely from the differences in measurement methods between the various systems.

Dr. Neels's screens comparing automated TPF and TPH with FHP inappropriately exploits these differences. Since the respective measurement systems are independent, the screens cannot discern anomalies that are byproducts of the FHP conversion process from actual data collection errors. Nor did Dr. Neels make any effort to examine the anomalies for evidence of

actual data collection errors. Response to USPS/UPS-T1-7, Tr. 23/8473-8474. Dr. Neels turns to the conceptual definition of TPF (and TPH) in terms of FHP and subsequent handlings as a defense of his methods (response to USPS/UPS-T1-7(e), Tr. 23/8474), but this is not availing. As Prof. Roberts correctly notes, because FHP and TPF are measured independently, the conceptual definition will not hold as an inequality. Response to USPS/OCA-T1-19, Tr. 23/8326. Accordingly, some "anomalies" will arise solely due to the FHP weight conversions.

Screening with FHP is particularly inappropriate as a basis for criticism of the Postal Service models, which use the machine-counted TPF where available. Since the FHP and machine counts are independently determined, there is no prior reason to believe that FHP conversion errors would indicate errors in the automated systems that provide automated TPF and TPH counts. Indeed, avoiding the need to account for the effects of weight conversion on FHP is a significant reason for preferring the machine-counted workload data.

III.B.1. Some of Dr. Neels's Screening Criteria, Perversely, Identify More Anomalies When the Errors are Relatively Small.

Dr. Neels's TPH-FHP and TPF-FHP screens also have the perverse feature that in operations where TPH and FHP should be relatively close, relatively small errors in the FHP conversions will trigger the identification of more "anomalies" by Dr. Neels's criteria. Certain operations, such as those in the OCR cost pool, generally are the first sorting operations for pieces processed therein. Furthermore, the successfully processed pieces are generally handled

in other operations for subsequent handlings (*i.e.*, D/BCS operations, in the case of successfully OCR'ed letters). Since true TPH and FHP are nearly equal in such operations, relatively small overestimates of FHP in the conversion process will trigger "anomalies" of the sort identified by Dr. Neels.

Dr. Neels agrees that operations where actual TPH and actual FHP are close will be more susceptible to the identification of anomalies according to his screens. Tr. 23/8556. In fact, in the limit where true TPH and FHP were equal, and the conversion factors were accurate on average, Dr. Neels admits that his screening procedures would identify half the observations as anomalous. Response to USPS/UPS-T1-36, Tr. 23/8528. Indeed, Dr. Neels's screen identifies the most FHP-TPH anomalies in the OCR operation, some 38 percent of observations. UPS-T-1 at 18. Particularly given the use of instrumental variables estimation procedures that are robust to measurement error from the FHP conversion process, there is no reason at all to consider those anomalies to constitute observations that are unusable for estimation. This screen is clearly inconsistent with Dr. Neels's approach to minor data anomalies in Docket No. R97-1.

III.B.2. Some of Dr. Neels's Screening Criteria Defeat the Purpose of Cost Pools.

The MODS cost pools serve two significant and long-acknowledged purposes. They consolidate operations into analytically significant groups of workhours and/or volumes. They also are robust to certain forms of data errors at finer levels of operation disaggregation. When data are booked to operation

A, but should have been booked to operation B in the same cost pool, the operation-level data may be in error but the cost pool total will be correct. Dr. Neels agrees. Response to USPS/UPS-T1-11(d), Tr. 23/8484-5. Indeed, in Docket No. R2000-1, Dr. Neels even had attributed some differences in results between his models and the Postal Service models to elimination of certain errors in parcel volumes by aggregation of certain cost pools' data. Response to USPS/UPS-T1-38, Tr. 23/8533.

In my direct testimony, I had described anomalies in the data for certain barcode sorting operations that were detected when disaggregated data were provided to Prof. Roberts in Docket No. R2005-1. It had appeared that data were being incorrectly booked between MPBCS and DBCS operations in similar schemes, but the totals over the incoming and outgoing D/BCS cost pool groups were correct. USPS-T-12 at 50, lines 16-21. These issues particularly affected the FHP data. In discussions with MODS staff at Postal Service Headquarters, I was able to determine that incorrectly booked FHP withdrawal transactions were responsible for the observed errors, which indicates that the D/BCS aggregates would be correct.

Indeed, Dr. Neels confirmed that rates of negative-FHP errors for D/BCS operations are trivial at the level of the Postal Service cost pools. Response to USPS/UPS-T1-9(c), Tr. 23/8480-81. Dr. Neels actually uses the FHP at even higher levels of aggregation. Despite the indications that the aggregates were correct, Dr. Neels screened the data at sub-cost pool levels that would show more erroneous observations. This is inconsistent with the principle that

anomalous but valid observations should be employed for analysis. Again, it is hard to see a purpose to Dr. Neels's screening procedures other than to make the MODS data look bad.

III.B.3. The AP and Weekly Productivity Screens Also Tend to Identify Relatively Small Data Errors.

Dr. Neels also identifies as anomalous or erroneous observations in which productivity screens fail for some of accounting period or weekly data that make up the quarterly observations. UPS-T-1 at 18, 20 (Tables 4 and 5). The observation that certain errors may be "masked" in aggregated data was a bailiwick of Dr. Neels's in Docket No. R2000-1, and he effectively repeats the criticism here, paying only lip service[6] to the underlying theoretical point that errors too small to be noticed are unlikely to have material effects on the estimation procedures. USPS-T-12 at 65, lines 6-17.

Screening at the weekly frequency, not surprisingly, vastly increases the amount of data that could be in error: behind the 10,304 quarterly observations in the USPS-LR-L-56 data set are 133,952 weekly observations. Given a thirteen-fold increase in observations, even relatively low error rates of a few percent in the weekly observations would lead to the elimination of most of the quarterly data, since it would be expected that most observations would contain at least one bad week. Even Dr. Neels's screening shows that not to be the case. Moreover, an error in weekly data that would, on average, constitute 1/13th of the

[6] "Not all errors are equally important." UPS-T-1 at 21.

quarterly observation would clearly tend to be small relative to an error in the quarterly or even AP data.

The main issue is whether the small errors have a material effect on the results. The effects, as I showed, are generally minor, particularly for the letter and flat operations that comprise the vast bulk of the costs under study, and show no evidence of significant bias in any direction.[7] USPS-T-12 at 97, 99. With the vast majority of the "masked" errors showing no signs of causing actual problems for the estimation procedures, Dr. Neels has again failed to demonstrate that the observations are unusable.

III.B.4. The Use of Robust Estimation Procedures Reduces the Need for Screening Based Solely on FHP Conversion Error.

Finally, changes in estimation methods since Dr. Neels's previous testimony make his data screening approach particularly excessive. Unlike the previous proceedings in which the Postal Service models used generalized least squares estimation procedures that may theoretically have been susceptible to errors-in-variables problems, the current models use instrumental variables procedures where indicated.[8] USPS-T-12 at 86 to 89 (line 4). In particular,

[7] Manual Priority is the notable exception, but the standard errors of the estimates also increase dramatically for the smaller samples, such that the resulting differences are not statistically significant. The analysis did not seek to locate observations that may have been serving as leverage points.

[8] Dr. Neels's swipe at Prof. Greene for purportedly failing to predict the effects of introducing instrumental variables estimation (Tr. 23/8512) is misdirected. The qualitative, and many of the quantitative, results of the Postal Service analysis are robust to the introduction of instrumental variables procedures. This is evident even in Prof. Roberts's Table 1, where the "USPS" results use

screening based solely on the presence of conversion errors in FHP is unwarranted, since the estimation procedures are intended to be robust to that source of errors-in-variables.

III.C. Errors and Omissions Make Dr. Neels's Analysis of the Postal Service Models Highly Misleading.

III.C.1. Dr. Neels Concedes the Robustness of the Postal Service Models' Main Results.

Perhaps intending damnation by faint praise, Dr. Neels opens his discussion of the Postal Service variability models noting:

> While this latest study [in USPS-T-12] incorporates more recent data, it is based on the same conceptual framework used in the past. Not surprisingly, it has produced substantially the same results. (UPS-T-1 at 9, lines 3-5.)

When Prof. Bradley initiated the econometric study of mail processing volume-variability in Docket No. R97-1, the robustness of the Postal Service models was very much in question. The robustness of the models has previously been assessed in very inappropriate ways, for instance by comparing results of statistically rejected econometric estimation approaches with those from appropriate models.

Assuming *arguendo* that Dr. Neels were correct as to the unreliability of the MODS data and the major features of mail sorting operations that are supposedly ignored by the Postal Service models, it would be fair to wonder just how the models manage to pull off the trick of consistency. In fact, the problem

---

instrumental variables estimation even where the Postal Service models do not, and yield similar results overall. OCA-T-1 at 13.

is with Dr. Neels's premises and not the Postal Service models or results. As discussed above, he overstates data quality issues with MODS. As I discuss below, he likewise mischaracterizes the purported failings of the Postal Service models by employing analyses that are biased towards showings of instability.

III.C.2. Dr. Neels's Split Sample Analyses Fail to Demonstrate Truly Significant Problems with the Postal Service Results.

Dr. Neels conducts several specification tests purporting to show instability of the Postal Service models over subsamples of regression observations. He interprets the rejection of pooling as indicating flaws in the Postal Service models. UPS-T-1 at 31-36. In this case, his interpretation of the results is not facially wrong, but it is highly misleading.

First, Dr. Neels tests the full set of slope coefficients, rather than the specific coefficients that enter the variability coefficients. Response to USPS/UPS-T1-13; Tr. 23/8487-8489. As a result, while he bemoans the need to estimate "nuisance parameters" in microeconometric models (response to USPS/UPS-T1-15; Tr. 23/8492), he is happy to conduct tests that exploit them for his ends. His testing procedures allow there to be no significant differences in the parameters that determine the variabilities, but still lead to a "rejection" based on differences in other parameters.

More importantly, Dr. Neels's critique conspicuously fails to show that the rejection has significant effects on the quantities of interest, elasticity estimates applicable to cost pools or components applicable to alternative models. Indeed, in the code for his tests of operations with translog models, Dr. Neels did not

even compute the elasticities. For example, UPS-WP-1, program WP_Chow_Big vs Rest.do. This is an especially significant omission for the translog form, where individual coefficients generally do not have economic interpretations, so computing elasticities and like quantities is central to interpreting the model results. Also, if models estimated on subsamples are to be used, the subsample results need to be combined, for instance, through a weighted averaging procedure. Response to USPS/UPS-T1-13, Tr. 23/8487-8489.) Dr. Neels confirmed the combined split-sample and weighted results in Table 3, below, derived from the models Dr. Neels estimated for Table 11 of UPS-T-1. (Id.)

Table 3. Combined Split-sample results Versus Postal Service BY 2005 Models, Small vs. Large Plants

Cost Pool	Variability, "Big Plants" Sub-sample	"Big Plants" Share of FY05 Hours	Variability, "Small Plants" Sub-sample	"Small" Share of FY05 Hours	Weighted Average Variability, Cost Pool	Neels replication of USPS BY05 Variabilities	Difference
OCR	0.72 (0.06)	0.87	0.91 (0.06)	0.13	0.74 (0.05)	0.78 (0.05)	-0.04
FSM 1000	0.74 (0.03)	0.79	0.67 (0.04)	0.21	0.73 (0.02)	0.72 (0.03)	0.01
SPBS	0.87 (0.04)	0.92	0.92 (0.07)	0.08	0.87 (0.04)	0.87 (0.05)	0.00
Incoming D/BCS	0.86 (0.06)	0.83	0.69 (0.07)	0.17	0.83 (0.05)	0.82 (0.07)	0.01
Outgoing D/BCS	1.00 (0.06)	0.89	1.08 (0.05)	0.11	1.01 (0.05)	1.06 (0.06)	-0.05
Weighted Average					0.85	0.86	-0.01

Thus, while Dr. Neels's testing procedure was able to generate "rejections" for his Table 11, the practical effect of splitting the sample and combining the results is negligible. As a result, he fails to show the existence of a material problem with the Postal Service models.

A similar pattern holds for Dr. Neels's Table 13 analysis, for which he split the sample between plants with increasing volumes and plants with decreasing volumes. Again, Dr. Neels's "rejection" fails to translate into material changes in the variability results.

Table 4. Combined Split-sample results Versus Postal Service BY 2005 Models, "Growing" vs. Other Plants

	Split-Sample Weighted Average			USPS BY2005		Differences
Cost Pool	V.V.s	S.E.s		V.V.s	S.E.s	
OCR	0.84	0.06		0.78	0.05	-0.05
BCS-Incoming	0.84	0.07		0.82	0.07	-0.02
BCS-Outgoing	1.04	0.05		1.06	0.06	0.02
FSM1000	0.71	0.03		0.72	0.03	0.01
SPBS	0.86	0.05		0.87	0.05	0.00

II.C.3. Dr. Neels's Fixed-Effects Analysis Inappropriately Fails to Account for the Statistical Properties of the Estimates.

Dr. Neels presents the results from an analysis of the estimated fixed effects coefficients from the Postal Service models which, he claims, demonstrates an "implausible" pattern of relative productivity differences among sites. He marvels at the results and wonders if there might not be something "deeply wrong" with the models. UPS-T-1 at 36, line 7, to 38, line 2. When asked to produce similar results for his alternative model from Section 6 of UPS-T-1, he reported comparably large ranges of productivities. Response to

USPS/UPS-T1-29; Tr. 23/8511-8514. A reasonable question would be whether there is something deeply wrong with all of the models, or if a more innocent explanation is available. A correct interpretation indicates the latter.

The key consideration is that the estimated fixed effects are statistical estimates and thus are subject to sampling-type variation. This fact points to one major problem with Dr. Neels's analysis, which is that he only examined the extreme (maximum and minimum) values of the effects. Looking at enough estimated values, such as the 200 to 300 sites' effects in the Postal Service models, one would *expect* to see some number of large deviations from the "true" values of the coefficients solely due to random variation. Not surprisingly, it is therefore rare to summarize a collection of data using only the extreme values.

Dr. Neels's calculations also contribute to the wide measured range. Since the dependent variable of the labor demand equations is the natural logarithm of hours, he exponentiates the estimated fixed effects to obtain the effects in the level of workhours. This is conceptually appropriate. However, Dr. Neels apparently fails to consider the effects of the exponentiation and standardizing the results based on extreme values on the sampling variability of his reported results. These steps have the effect of magnifying large deviations from the "true" values. This is highly apparent in Table 5, which shows that the standard errors for Dr. Neels's productivity ratios are high relative to the ratios themselves.

Table 5. Standard Errors of Neels Productivity Ratios

Cost Pool	Median	Max	SE(Max)	Standard Difference

Manual Letter	2.07	5.03	1.07	2.76
Manual Flats	1.95	4.95	0.89	3.37

What Dr. Neels has discovered is not a fundamental flaw in the models—the Postal Service models or, for that matter, his own. Rather, he has set up an analysis in which the random variation in the productivity factors is large relative to the productivity differences he is purporting to measure. Thus, his results are driven in no small part by the noise in the coefficient estimates. He would, in fact, have to be "unlucky" not to see a wide range of effects under the circumstances. Consequently, his analysis does not provide a legitimate critique of the fixed-effects models.

III.C.4. Dr. Neels's Analysis of Instrument Relevance has No Statistical Content At All and Must Be Rejected.

Dr. Neels purports to assess the relevance of the instrumental variables employed in the Postal Service models by examining a "partial R-squared" statistic from a regression of the instrumented output variables on the excluded instruments. UPS-T-1 at 30. While he loosely motivates the analysis by citing to a well-known paper[9] on instrumental variables analysis with "weak instruments," the referenced paper does not actually propose a test based on the "partial R-squared." Dr. Neels admitted that the criteria he used to determine whether "weak instruments" might lead to bias in the Postal Service models were not derived from any formal statistical test at all. Response to USPS/UPS-T1-12, Tr. 23/8486. As such, it is impossible to evaluate whether the criteria Dr. Neels

[9] Douglas Staiger and James H. Stock, "Instrumental Variables Regression with Weak Instruments," Econometrica 65(3), 557-586.

employed are valid, and the Commission should not rely on Dr. Neels's pure judgment.

Moreover, Dr. Neels's judgment did not seem to account for significant features of the paper he cited as they apply to the Postal Service analysis. The Staiger and Stock paper shows, "In contrast to TSLS two-stage least squares], [the limited information maximum likelihood, or LIML, estimator] rapidly becomes median unbiased" as the instruments become stronger.[10] They described the result as supporting previous observations of small bias in LIML in more narrowly defined circumstances. This is part of the reason why the Postal Service models employ LIML. Dr. Neels agrees that Staiger and Stock distinguished the performance of LIML and TSLS (response to USPS/UPS-T1-12, Tr. 23/8486), but it is impossible to determine whether or how he might have modified his judgment accordingly. Finally, the nature of weak instrument bias is such that were the LIML estimator exhibiting the bias Dr. Neels purports to show, the LIML results should be much closer to the ordinary least squares results than they actually are. USPS-T-12 at 87. Again, Dr. Neels adduces no real problem with the Postal Service instrumental variables models.

III.C.5. Dr. Neels's Analysis of Equipment Deployments Does Not Demonstrate a Simultaneity Problem.

As part of a discussion that of changes over time to Postal Service operations, Dr. Neels presents results of a logit model that purports to show a

[10] Staiger and Stock, moreover, modeled instrument "weakness" as nearly zero correlations between the instruments and exogenous variables.

relationship between volumes and equipment deployments. UPS-T-1 at 42-43. Asked to describe the structural model of equipment deployment underlying his analysis, he deprecated his own analysis saying it "is not intended to describe a fully-articulated model of technology deployment" (response to USPS/UPS-T1-34(a), Tr. 23/8524), which is an understatement.

It does not take a fully-articulated model of technology deployment to understand that Postal Service managers do not simultaneously juggle the mail they need to process and the equipment needed to process it. Rather, equipment deployments are determined well in advance of the actual operational use of the equipment. My understanding is that achieving a threshold return on the Postal Service's investment is among the criteria that interacts with deployment decisions, and obtaining that return with the mail processing equipment Dr. Neels considers involves achieving marginal cost reductions applied to sufficiently large volumes of mail. Observing large, expensive, high-capacity pieces of sorting machinery at large plants does not indicate that capital and volumes are simultaneously determined, it simply suggests that Postal Service planners are rational in deciding where to put equipment.

IV. Dr. Neels's Alternative Model, Correctly Applied, Supports the Postal Service Variability Results.

Dr. Neels advances the alternative model he presents in Section 6 of UPS-T-1 (pages 49-54) as a "plant-level" alternative to the Postal Service models. As implemented it is not truly a "plant-level" model, but instead an aggregated model of the sorting operations covered by the Postal Service analysis. *Id*. at 49, lines 14-15.

As I discuss below, as a model of sorting operations, Dr. Neels's alternative model is highly inappropriate. It abandons, without good cause, the shape-based processing structure that characterizes both actual Postal Service operations and the economic frameworks underlying the Postal Service and OCA models. The aggregated structure Dr. Neels proposes totally fails to reflect features of operations, such as cost differences between manual and automated mailstreams, that Dr. Neels admits are characteristics of the operations. Dr. Neels's operational motivations for his approach sorting operations do not stand up, and he failed to estimate his models in ways that would have provided direct evidence to support his contentions. In fact, the evidence suggests that the cross-shape and cross-operation effects he describe do not exist to any significant extent.

As a model of non-sorting operations, Dr. Neels's alternative model is potentially more useful. I describe below how its structure generally resembles models of "allied labor" operations originally advanced by the Postal Service in Docket No. R97-1. It may also provide an empirical method to estimate overall

variability for mail processing plants, though the need to de-average sorting operations would remain. I show below that true "plant-level" versions of Dr. Neels's models generally support the Postal Service variability levels.

IV.A. Dr. Neels's Alternative Model Is Inappropriate for Sorting Operations

IV.A.1. Dr. Neels's Account of Sorting Operation Activities Does Not Suggest Significant Cross-Operation Effects.

Dr. Neels presented an account of sorting operation activities that purports to provide support for his contention that cross-operation and even cross-shape effects must be taken into account in modeling sorting operations. Response to USPS/UPS-T1-2, Tr. 23/8457-8460. Correctly interpreted, Dr. Neels's account shows only minor cross-effects. Indeed, I tested for the presence of certain cross-operation effects and presented results in USPS-T-12 that show them to be insignificant for automated letter and flat operations. USPS-T-12 at 93-95.

Discussing the "runtime" activity, Dr. Neels cites the possibility of piece handlings in one cost pool affecting the composition of mail in another, and references the past inclusion of the "manual ratio" variable in the Postal Service models. Dr. Neels admits that this effect, to the extent it is present, primarily operates "within the same shape-based mailstream." Response to USPS/UPS-T1-32(c), Tr. 23/8520. The citation of the mail composition effect that the manual ratio sought to capture does not directly imply cross-operation effects, though. The operational effect to which Dr. Neels alludes stemmed from changes in mail composition within the automated operations as the Postal Service's automation program advanced. The issue was not that manual volumes had a direct effect

on automated operations, but rather that the automation operations themselves were expanded to process “harder” mail. The practical utility of the manual ratio was that it proxied composition changes within both the manual and automated operations. While Dr. Neels’s interpretation is overly literal, it is subject to the testing showing the absence of cross-operation effects, described above.

With respect to container handling costs in sorting operations (“quasi-allied labor” in the USPS-T-12 terminology), Dr. Neels raises the prospect that congestion in the plant might affect the costs. While this might appear to be a potential cross-operation effect at first glance, Dr. Neels fails to consider which traffic would have the largest effect on these particular handlings.[11] He also admits that the volume of mail being moved should be the primary determining factor for the contatiner handling labor requirements. Response to USPS/UPS-T1-32(e), Tr. 23/8521. In Postal Service plants, the equipment associated with the mail processing cost pools are located together, and mail slated for subsequent processing (as opposed to mail to be dispatched from the facility; this dispatching work is generally carried out in LDC 17 operations) is typically staged close to the equipment. Thus, much or most of the traffic for Dr. Neels’s story will come from the same operation. Again, the presence of effects from other operations is testable.

[11] Dr. Neels indicates having only visited one Postal Service mail processing facility, in the course of Docket No. R97-1. Response to USPS/UPS-T1-32(a), Tr. 23/8520.

IV.A.2. Dr. Neels Fails to Demonstrate Cross-Operation and/or Cross-Shape Effects, While Inappropriately Failing to Represent Key Patterns of Cost Causation.

Despite claiming the presence of significant cross-operation and cross-shape effects, Dr. Neels does not actually demonstrate that such effects exist. Dr. Neels confirms that, in examining shape-level alternatives to his aggregated model, he did not investigate any models directly incorporating cross-shape or cross-operation effects. Response to USPS/UPS-T1-5(b), Tr. 23/8468-8469; response to USPS/UPS-T1-32(d), Tr. 23/8521. Given the prominence Dr. Neels affords those effects in criticizing the Postal Service (and, by extension, OCA) models, not estimating models that would directly confirm the existence of the effects seems to be a major omission.

Meanwhile while Dr. Neels agrees that shape and automation compatibility determine how mail is handled and have effects on costs (responses to USPS/UPS-T1-15(e)-(g), 16-19; Tr. 23/8491-8493), yet the aggregation in his models leads him not to account for those factors at all. This contrasts with Prof. Roberts's correct suggestion that representing such cost differences between and within mailstreams motivates models that explicitly account for them. Responses to USPS/OCA-T1-1-2, Tr. 23/8287-8288.

Dr. Neels defines his "shape"-level outputs in a manner that somewhat obscures the shape-related mailstreams. He assigns Priority Mail handlings back to shape, and includes all SPBS handlings, including handlings of flat-

shape mail bundles, to the parcel category.[12] This treatment would have negligible effect on the letter FHP output in Dr. Neels's models, since Priority Mail letters are a relatively small category. Since the details of the non-letter treatment would not matter much in principle, I therefore estimated a model using workhours for letter-shape operations incorporating cross-shape effects using Dr. Neels's FHP measures. Principal results are presented in Table 6, below. As shown in the table, the sum of the "flats" and "parcels" elasticities is zero. This contrasts with a combined elasticity of 0.15 (significantly different from zero) in Dr. Neels's aggregated model. The results show that the plant model is not directly applicable to the letter-shape operations, and that, whatever the theoretical merits of Dr. Neels's claims of cross-shape effects, there is no evidence that they are actually present for letter-shape sorting operations.

Table 6. Results from Letter-Shape Version of Neels Model

	"Plant-level"		Letter-Shape Operations	
Output	coef	se	coef	se
FHP Letters	0.88	0.04	0.98	0.05
FHP Flats	0.09	0.01	0.01	0.01
FHP Parcels	0.06	0.01	-0.01	0.01
SUM	1.03	0.04	0.98	0.04

[12] The issue with this treatment is that most Priority Mail flats processing is outside of "regular" flats operations.

IV.B. Applied to Allied Labor and Broader Plant Labor Hours, Dr. Neels's Model Gives Results Consistent With the Postal Service Methods

IV.B.1. As a Model of Allied Labor Operations, Dr. Neels's Model Is Similar To Models Previously Advanced By the Postal Service

As a model of allied labor operations, Dr. Neels's model generally resembles the models advanced by Prof. Bradley in Docket No. R97-1, which related workhours in allied labor cost pools to MODS sorting workloads representing several letter and flat processing mailstreams (Docket No. R97-1, USPS-T-14 at 37-38). Dr. Neels's model measures letter and flat workload as the respective shape-level first-handling pieces, and adds a third workload variable representing parcel and bundle handlings. UPS-T-1 at 50. While Prof. Bradley's models used the translog functional form, Dr. Neels's use of the log-linear functional form in conjunction with fixed effects/instrumental variables estimation methods is reasonable.

While some of Dr. Neels's data handling decisions are questionable – including his excessive application of screens designed to identify errors in automated TPF to the FHP data he employs, as discussed above, and allocating Priority Mail handlings to shapes and combining parcel and bundle handlings – the broad outlines of the model are acceptable for the reasons articulated by Dr Bradley in Docket No. R97-1.

In the analyses of allied labor and broader plant workhours, I employ Dr. Neels's FHP measures and the larger of his two samples to minimize methodological differences with Dr. Neels's reported models.

IV.B.2. Contrary to Dr. Neels's Claims, His Models Could Have Been Estimated with Allied Labor Workhours

Dr. Neels is incorrect in claiming that he lacked data suitable to estimate his model for allied labor operations. UPS-T-1 at 49. The data sets originally provided in USPS-LR-L-56 included workhours from the National Workhour Reporting System (NWRS) for LDC 17, the Labor Distribution Code to which the allied labor cost pools belong. USPS-LR-L-56 also provided NWRS workhours for LDC 13, which in addition to SPBS and LIPS operations includes APPS, mechanized sack sorting, mechanized tray sorting, and robotics operations. Using these data would have permitted Dr. Neels to extend his analysis to cover an additional $3.52 billion in BY 2005 mail processing cost.

Using the NWRS data from USPS-LR-L-56, I estimated Dr. Neels's models for LDC 17 allied labor as well as for the "plant-level" hours including the LDC 13 and 17 workhours not incorporated in Dr. Neels's testimony. The results are provided in Table 7, below. Neither version of Dr. Neels's model shows a significant difference in the total elasticity from the sorting operation composite employed in the Postal Service CRA.

Table 7. Results of Neels Mail Processing Model Applied to LDC 17 (Allied Labor) Workhours

Output	LDC 17	LDCs 11-14; 17	USPS BY 2005 Sorting Operation Average
Letter FHP	0.76 (0.08)	0.74 (0.07)	
Flat FHP	0.07 (0.01)	0.03 (0.01)	
Parcel FHP	0.01 (0.01)	0.05 (0.01)	
Total	0.84 (0.07)	0.82 (0.06)	0.85

IV.B.3. Dr. Neels's Model Also Shows Variabilities Comparable to the Postal Service CRA When Applied to Total "Function 1" Workhours

Dr. Neels states that his intent was to apply his model to total plant workhours, and his use of the aggregate of the sorting and cancellation operations' workhours was a compromise driven by data availability. UPS-T-1 at 49. As I noted in the previous section, Dr. Neels actually had failed to use data of significantly broader scope that had been available in the Postal Service's datasets all along, and analyzing those data with Dr. Neels's model supports the Postal Service's overall variabilities for plants.

UPS requested MODS and NWRS data at still-higher levels of aggregation, including the "Function 1" aggregate. Total "Function 1" workhours encompass, in addition to the hours analyzed in the previous section, supervisory workhours from LDC 10 and workhours in LDC 18 "miscellaneous" operations (e.g., Express Mail, Registry and Mail Processing Support). I estimated a "Function 1" model, using Dr. Neels's specification, with the "Function 1" MODS workhours provided in USPS-LR-L-190.[13] The costs associated with the supervisory workhours are included in Cost Segment 2, and have been a "piggybacked" cost component assumed to be volume-variable to the same extent as the associated mail processing labor costs in Cost Segment 3.1 (See USPS-LR-L-1).

[13] I used MODS rather than NWRS data for this analysis because the NWRS data frequency changed in FY 2004, when the Postal Service adopted the U.S. Government fiscal calendar. Prior to FY 2004, the MODS and NWRS data are very highly correlated (correlation coefficient = 0.99). Conceivably, a future analysis of these data could recast the data to a common frequency, similar to the methods used with the MODS data from USPS-LR-L-56.

Results from the model with total "Function 1" hours are presented in Table 8, below. While the total elasticity is slightly higher than that from the NWRS-based model reported in Table 7, above, the result nevertheless does not differ significantly from the sorting operation composite used in the Postal Service CRA, and the variability is significantly less than 100 percent.

Table 8. . Results of Neels Mail Processing Model Applied to Total Function 1 (MODS Plant) Workhours

Output	Neels Model, Function 1 Hours	USPS Sorting Operation Average
Letter FHP	0.82 (0.03)	
Flat FHP	0.03 (0.01)	
Parcel FHP	0.05 (0.01)	
Total	0.90 (0.03)	0.85

V. Prof. Roberts's Models are Potentially Workable, but Would Benefit from Significant Improvements

V.A. Prof. Roberts's Theoretical Framework Is Conceptually Valid, but Cannot Fully Specify the Empirical Models.

V.A.1. Prof. Roberts Fails to Articulate the Relationship Between His "Outputs" and the Costs He Purports to Explain.

While Dr. Neels's proffered operational explanations of various component activities of sorting operations do not provide much if any support for his theories of cross-operation and cross-shape cost effect, as discussed above, Prof. Roberts provides no explanation at all of the operational substance behind his characterization of sorting operations' outputs. (Response to USPS/OCA-T1-3d, Tr. 23/8291.)

Prof. Roberts's dismissal of the constituent activities as "narrow" is malapropos considering that runtime, the largest of the "narrow" activities, comprises more than three-fifths of the total letter and flat sorting labor Prof. Roberts studies, and more than three-quarters of the "direct labor" (i.e., costs excluding "overheads"). The other activities, while much smaller, nevertheless constitute large amounts of the labor usage Prof. Roberts is studying. As I explained in my direct testimony (USPS-T-12 at 27-29), and Prof. Roberts does not counter, there is a clear engineering relationship between labor usage for sorting operation activities and the total number of sorts, *i.e.*, TPF. At some point, there must be tangible resource usage effects not captured by TPF. As

discussed above with respect to Dr. Neels's account of the activities, effects outside of own-operation TPF are small.

V.A.2. A Correct Understanding of the Relationship Between TPF and FHP Indicates a Closer Relationship Between the OCA and Postal Service Models than Prof. Roberts Admits.

The relationship between sorting workhours and TPF does not suggest that the characterization of output in Prof. Roberts's theoretical framework is wrong. In fact, the Postal Service method likewise conceptually starts with an array of mail products differentiated by subclasses, presort level, required sort depth, and other cost-causing characteristics. (USPS-T-12 at 48, lines 14-21.) For reasons I have discussed at length, it is not practical to base an applied analysis on direct measurement of costs for the full array of products; Prof. Roberts seems to have discovered this in pursuit of his models with "additional outputs." The central empirical question is how to use the Postal Service's "operating plan" to usefully characterize sorting operations' outputs. Response to USPS/OCA-T1-40, Tr. 23/8374. This is where Prof. Roberts's models fall short.

Prof. Roberts's 2002 and March 2006 papers have not evinced a solid understanding of the distinctions between the MODS TPF (or TPH) and FHP workload measures. Prof. Roberts's tendency has been to conflate FHP with "plant volume" (not recognizing that a significant portion of "plant volume" bypasses outgoing and/or incoming sorting operations) while considering TPF and TPH to be measures of "capital input" (Roberts 2006 at 36).

In fact, TPF and FHP measure different aspects of sorting workload. FHP measures the number of pieces sorted, but not the amount of sorting work performed; TPF measures the number of sorts but not the unique pieces worked. The practical question is which of these measures better reflects the causes of labor usage in sorting operations—is it the number of sorts or the number of pieces that is more relevant? The engineering relationships point to the number of sorts. So, as witness Oronzio notes, automated TPF has been preferred by Postal Service operations precisely because it has a closer and more stable relationship with workhours than FHP, and is measured more accurately besides (USPS-RT-15). It is not theoretically impossible to measure sorts with FHP, but it makes little sense to do so when direct measures exist.

Some of the properties that make FHP a less desirable analytical tool from a management perspective, such as its inability to show the subsequent processing paths of pieces, also make it difficult to practically distinguish whether the OCA model really is capturing effects outside of the TPF-based Postal Service models. For example, Prof. Roberts's "additional outputs" models define BCS and OCR/ISS FHP as separate "outputs" for letter-sorting operations. Response to USPS/OCA-T1-45(e)-(f), Tr. 23/8381-8383. The flow of successfully OCR'ed pieces to BCS operations would lead one to expect that both "outputs" would have significant effects on BCS resource usage, which Prof. Roberts's models generally show. OCA-T-1 at 42.

However, it cannot be concluded that a BCS model using BCS TPF is therefore misspecified. The actual mechanism whereby the OCR FHP affect

BCS labor usage matters. If the BCS labor usage due to OCR FHP stems from the subsequent BCS sorts, then there is no problem with the TPF-based model, since the downflow from the OCR FHP is captured in BCS TPF. This is a key respect in which the Postal Service and OCA models are more closely related than Prof. Roberts is willing to admit.

Prof. Roberts also agrees that certain "outputs" may totally bypass particular operations – an important case is pieces inducted into manual operations bypassing the automated mailstream – in which case those outputs may not appear in the bypassed operations' labor demand functions. Tr. 23/8406-8407. Indeed, as will be shown below, Prof. Roberts's models show (if not very efficiently) no effect of manual FHP on automated operations' labor usage. Depending on the details of the specification, including manual FHP as "outputs" may lead a loss of efficiency, or it may lead to more serious errors if "outputs" with different marginal labor costs are forced to have a common effect on workhours.

V.B. Prof. Roberts's Updates to the Two-Output Models are Unjustified and Inconsistent With His Previous Methods.

The models Prof. Roberts presents in OCA-T-1 incorporate several significant changes in econometric methodology, mainly involving sample and instrumental variable selection. OCA-T-1 at 17-18, 27-31. Compared with the methods Prof. Roberts articulated in his 2002 and March 2006 papers, which he cites as background to his model-building exercise (OCA-T-1 at 3-5; response to USPS/OCA-T1-35, Tr. 23/8364), the current changes appear arbitrary and bear

justifications that are in conflict (if not flatly contradictory to) the previous work. Prof. Roberts also commits serious econometric errors in his treatment of seasonal variation for his analyses of workhours and piece handlings—errors which would have been completely avoidable had Prof. Roberts simply continued with the correct treatment from his earlier research.

V.B.1. Prof. Roberts's Sample Period is Unnecessarily Short, Particularly in Light of His Previous Justification of Much Longer Sample Periods Covering Equally Significant Operational Changes.

The major sample change in Prof. Roberts's OCA-T-1 models is to shorten the sample period to FY2002-FY2005, allowing up to 16 quarterly observations per site, versus the FY1999-FY2004 period from the March 2006 paper (50 percent longer) and the FY1994-FY 2000 period (75 percent longer) from the 2002 paper. Tr. 23/8419. Prof. Roberts's main explanation for the change is intended to deal with changes to flats processing due to AFSM deployment, though he also extends the treatment to letter operations that are not affected by flats processing changes. This treatment is questionable for flats, barely justified for letters, and inconsistent with Prof. Roberts's previous work.

There is no doubt that the AFSM 100 has represented a major advancement for automated flats processing, both by expanding automation capacity and increasing automation productivity. However, Prof. Roberts's claim that it represents a change such that pre-AFSM data from *non-AFSM* flats operations cannot be used—in contrast to his March 2006 work—is not well-justified. The transition from the FSM 881 to the AFSM 100 should have been

well-known to Prof. Roberts as of his March 2006 paper. The AFSM 100 has, of course, greatly increased the capacity and productivity of automated flat sorting equipment, and so has permitted automated processing of significant quantities of flat-shape mail that formerly had been processed manually. Of course, relieving automation capacity constraints with much more productive equipment should result in a reduction in measured marginal flat sorting costs.

The main issue, though, is that Prof. Roberts's results do not show the effect he claims as motivation. The reduction in sample size causes the standard errors of the estimates to rise, especially for the FSM 1000 elasticities, such that even the large change in the FSM 1000 point estimate is not statistically significant. As a matter of statistical logic, it is inappropriate to introduce a higher-variance estimation procedure, then to argue that there has been a structural change from statistically insignificant differences involving the less-precise estimates.

Moreover, while the changes to operations are substantial, they are not obviously more significant than the changes to letter-shape operations over the time period of Prof. Roberts's 2002 analysis. Prof. Roberts extensively discussed significant changes to letter operations over his sample period—which included DBCS deployment, LSM retirement, remote barcoding, and significant expansion of delivery point sequencing—and concluded, for the most part reasonably, that including control variables for technology mix in his models was a sufficient solution. Roberts 2002 at 19-21; see also Tr. 23/8423-8424. In contrast, the relatively minor changes to letter-shape sorting operations that Prof.

Roberts cites in justification of the shorter sample period for letter sorting operations do not warrant the change in methodology.

V.B.2. Prof. Roberts's Approach to Seasonal Dummy Variables Is Neither Operationally nor Econometrically Justified, and is Contrary to His Correct Earlier Treatment

Prof. Roberts introduces a new treatment of quarterly dummy variables in OCA-T-1—excluding them as explanatory variables from the labor demand equations and instead using them as instrumental variables (OCA-T-1 at 28-31)—that is inappropriate both operationally and statistically, and also represents an unwarranted about-face from his previous methods. The change is particularly pronounced from his 2002 paper, in which Prof. Roberts's correctly included quarterly dummy variables on similar grounds to those advanced by the Postal Service Roberts 2002 at 23-24.

V.B.2.a. Contrary to Prof. Roberts's Claims, the Postal Service Operational Testimony Did Identify Seasonal Factors Affecting Workhours.

Prof. Roberts claims not to have included seasonal dummy variables in his current models because he could not locate a discussion that would justify their inclusion in the Postal Service's operations testimony. Roberts 2006 at 59. In fact, accompanying the Postal Service's introduction of MODS-based costing methods in Docket No. R97-1, witness Moden described seasonal factors affecting Postal Service operations:

> …[T]he volume and characteristics of the mail vary significantly with the mailing seasons. The largest seasonal effect is certainly the

> holiday mail each November and December. The beneficial effect on productivity of a greater mail volume is overwhelmed by detrimental changes in mail characteristics and the temporary surge in staffing requirements. There is a large increase in letters and packages with illegible handwriting and incomplete addresses. Many of the letters use colored envelopes which are difficult for the OCRs and many of the packages are poorly wrapped. To process the workload in manual operations, temporary clerks are needed who are significantly less productive than the regular staff. Docket No. R97-1, USPS-T-4 at 20.

In short, witness Moden described seasonal variations both in staffing patterns and in mail characteristics. So Prof. Roberts is not justified by the absence of operational discussion. Indeed, Prof. Roberts had argued similarly in favor of including quarterly dummy variables in his 2002 paper:

> 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. Roberts 2002 at 24.

In Prof. Roberts's March 2006 paper, his main results did not incorporate quarterly dummy variables to control for seasonal factors, but he indicated that he had estimated models including such seasonal controls. Roberts 2006 at 59.

In his present testimony, in contrast, Prof. Roberts claims never to have incorporated such factors in his models. Tr. 23/8412. Prof. Roberts seems to have discovered the Postal Service's operational explanation for seasonal controls and subsequently forgotten it.

V.B.2.b. Prof. Roberts's Conclusion that Failure of the Quarterly Dummy Variables as Instruments is a "Spurious" Result is Unwarranted.

In response to USPS/OCA-T1-12 (Tr. 23/8309), Prof. Roberts showed that his models largely fail "overidentifying restrictions" tests when the quarterly dummies are used as instruments, and largely pass when the quarterly dummies are not so used. While he reaches the (correct) conclusion that the quarterly dummies are inappropriate, he nevertheless judges the result to be "spurious." Given the prior reasons, including those previously expressed by Prof. Roberts himself, Prof. Roberts's assessment is not well-founded. He had ample reason to believe that the correlations among the variables were not coincidental.

Prof. Roberts's conclusion is not justified by past evidence that the quarterly dummy variables were irrelevant in the labor demand models. Instead, Prof. Roberts admits that certain of the coefficients on the quarterly dummy variables were statistically significant in his 2002 analysis, but denies that the statistical significance of the effects justifies his former inclusion of those variables. Tr. 23/8414-8415. Prof. Roberts's argument might be admissible if he

lacked a prior reason for the variables' inclusion, but that is not the case here. Excluding variables that are both significant and which are expected theoretically to play a role is highly unusual.

Moreover, regression theory indicates that Prof. Roberts's observation that the quarterly dummies are highly correlated with volumes makes it less appropriate to exclude them from the models, not more. A basic result for multivariate regression models is that given the model:

$$y = b_1 X_1 + b_2 X_2 + \varepsilon ,$$

an estimate of b1 from a regression excluding X_2 will be biased and inconsistent unless X_1 and X_2 are orthogonal (i.e., "uncorrelated").[14] In this case, with the quarterly dummies playing the role of X_2, Prof. Roberts admits that the conditions for excluding them from the model – the absence of correlation with X_1 (specifically, the included MODS workload(s) – do not hold.

Prof. Roberts's explanation that he wishes to make use of (more) seasonal variation in the FHP variables to measure the labor demand elasticities does not remedy the situation, but rather suggests what is likely to be a biased estimation method. In light of the *a priori* operational reason for accounting for seasonal variation in workhours, and the statistical evidence that such effects actually exist, Prof. Roberts's claim that the rejection of the quarterly dummy variables as excluded instruments is "spurious," (USPS/OCA-T1-12, Tr. 23/8309) is without merit.

[14] See, e.g., Peter Schmidt, *Econometrics*, New York: Marcel Dekker, 1976, pages 39-40.

V.C. The Postal Service's Updates of Prof. Roberts's March 2006 Models is Superior to the Prof. Roberts's Versions in OCA-T-1.

Prof. Roberts's March 2006 paper presented results for a new two-output model (with incoming and outgoing FHP by shape as the outputs) which Prof. Roberts estimated with Docket No. R2005-1 data, which extended through FY 2004. Given the likelihood of interest in an update from the Commission, the OCA, and other interested parties, as well as the Postal Service's own interest in understanding Prof. Roberts's models, I updated Prof. Roberts's work with relatively minor changes to accommodate more recent data, BY 2005 cost pool changes, and a partial response to capital measurement issues Prof. Roberts raised. USPS-T-12 at 101-104. For reasons discussed below, should the Commission adopt a version of Prof. Roberts's two-output model, it should employ the updated version of Prof. Roberts's model presented in USPS-T-12.

V.C.1. Prof. Roberts's Explanations for Rejecting the Postal Service Update are Incoherent.

While OCA-T-1 did not mention the USPS-T-12 update of Prof. Roberts's March 2006 models, Prof. Roberts was asked in USPS/OCA-T1-12(a) whether he had considered the model and, if so, the basis for rejection. Tr. 23/8306-8308. Prof. Roberts stated that he had, in fact, considered the model, and discussed four substantive points that apparently were the basis of his rejection. However, Prof. Roberts's explanation does not support rejection of the Postal Service updates, as the factors Prof. Roberts raises are not actually defects of the USPS-T-12 update. These were:

- *Addition of FY 2005 data to the March 2006 sample*. Prof. Roberts observes that this was a "fairly small change" yielding results "very similar" to those in his paper. As I discuss above, the FY 1999-FY 2005 sample period is appropriate. It would be extremely unusual to consider the failure of a new econometric model to "break" from the addition of more data to be a defect.
- *Implementation of updated capital data to mitigate capital timing issues raised in the March 2006 paper*. Again, partly mitigating the problem identified by Prof. Roberts is not a defect. Prof. Roberts seems to have amplified the importance of the anomaly, which appears to be a fairly straightforward result of reporting lags in the PEAS system. I have been investigating solutions to the problem that involve collecting equipment deployment schedules to identify the proper timing, and expect the problem should be amenable to a full solution. In any event, the limitations of the existing data did not prevent Prof. Roberts from recommending models that use it.
- *Use of FY 2005 weights to combine elasticities*. Once again, this is not ad defect, and Prof. Roberts adopted this method in his recommended models.
- *Implementation of incoming and outgoing D/BCS cost pools*. Prof. Roberts suggests that the change is not "well justified" but stops short of

claiming the approach to be incorrect.[15] Prof. Roberts suggests that the "theoretical model" would suggest splitting all operations, which is not unreasonable in principle. Prof. Roberts's argument has a 'make the perfect the enemy of the good' flavor in suggesting that it would be inappropriate to disaggregate two large cost pools with significant costs on both the incoming and outgoing sides if all cost pools are not treated the same. Regardless of the merits of the change, updated results using the BY 2004 cost pool structure incorporated in Prof. Roberts's March 2006 paper were also provided (USPS-T-12 at 127), so Prof. Roberts faced no obstacle to rejecting the change if he so desired. Reverting to the earlier cost pool structure would not have substantially altered the results.

Thus, Prof. Roberts's reasons were primarily reasons to accept, not to reject, the USPS-T-12 study; given the possibility of controversy over the BCS cost pool issue, I provided results without the change as well. As I discuss in the following sections, Prof. Roberts's updates do not constitute an improvement.

V.C.2. The Postal Service Update of Prof. Roberts's March 2006 Model Yields More Robust and Plausible Results, Including by Prof. Roberts's Own Explanations.

Prof. Roberts notes that the results from the USPS-T-12 update are similar to those in his March 2006 paper. Response to USPS/OCA-T1-12, op. cit. Indeed, I demonstrated the results from the sample update on a constant-

[15] Prof. Roberts also cites the related change to the AFSM 100 cost pool in the Postal Service models; that had not been implemented in the update of his flat-shape models.

methodology basis and the results are very stable from that perspective as well. USPS-T-12 at 104.

The same cannot be said for Prof. Roberts's updates. The changes in elasticities from the March 2006 paper are driven by very high measured elasticities for the manual letters and FSM 1000 cost pools. In manual letters, previous versions of Prof. Roberts's models had not unreasonably provided elasticity estimates statistically close to unity. Roberts 2002 at 102; USPS-T-12 at 127. Prof. Roberts suggests that diversions of volumes to manual operations may account for the present 152 percent variability. Response to USPS/OCA-T1-8(b), Tr. 23/8300, but witness Oronzio's testimony shows that Prof. Roberts's explanation is not consistent with Postal Service operating practices. USPS-RT-15 at 10-12.

In case of FSM 1000, the standard error of the estimated FSM 1000 elasticity also increased markedly, and the change derives specifically from the outgoing FHP elasticity. Response to USPS/OCA-T1-16, Tr. 23/8321-8322. In this case, the change is actually inconsistent with Prof. Roberts's own explanation for elasticities that derived mostly from incoming flats FHP from the March 2006 paper, in which he noted that the "volume of outgoing flats is small relative to the volume of incoming flats and… it is basically hard to detect any systematic relationship between this category of FHP and manhours [sic]. Roberts 2006 at 49. The results of the USPS-T-12 update are consistent with Prof. Roberts's correct explanation for the pattern of his previous results.

Both results appear to be outliers of sorts, and their lack of operational basis suggests strongly that the changes that led to them, particularly the much shorter sample period, have not actually improved Prof. Roberts's results. In contrast, there are no such anomalies in either Prof. Roberts's March 2006 results or the Postal Service update. Accordingly, the latter should be preferred.

V.D. Prof. Roberts's "Additional Outputs" Models Point the Way To Improvements, But Prof. Roberts's Implementation is Faulty.

As a characterization of sorting operations, Dr. Neels's alternative model is badly flawed by incorporating cross-shape relationships for which there is no evidence, while failing to allow for features such as cost differences between the automation and manual mailstreams that even Dr. Neels agrees exist. Response to USPS/UPS-T1-16-19, Tr. 23/8494-8497. Prof. Roberts, in contrast, correctly views it as potentially important to represent systematic cost differences between various mailstreams in the labor demand models (see, e.g., OCA-T-1 at 20; Response to USPS/OCA-T1-1, Tr. 23/8287). Indeed, many of the changes Prof. Roberts has pursued from his original 2002 model have been to that effect, including the models with "additional outputs" he examined for OCA-T-1. (OCA-T-1 at 20-24.) While a step in the right direction conceptually, Prof. Roberts's implementation is lacking. In particular, while Prof. Roberts cites to the "additional outputs" models in the course of agreeing that manual and automated mailstreams have significantly different costs (response to USPS/OCA-T1-1-2, op. cit.), his models do not actually distinguish the manual and automated mailstreams.

V.D.1. Prof. Roberts Inappropriately Includes Automation Volumes with "Nonautomation" Letters.

The signal failing of Prof. Roberts's efforts to disaggreagate FHP by other output characteristics is that his "nonautomation" output actually combines significant amounts of automation-compatible FHP for non-prebarcoded pieces with manual FHP. Response to OCA/USPS-T1-45, Tr. 23/8381-8383. Indeed, as shown in my response to TW/USPS-T11-1(b)-(c) (Tr. 10/2568; 2573), manual FHP are a minority of Prof. Roberts's "nonautomation" category. This would make distinguishing the costs for "true" nonautomation mail from automation mail difficult.

A correct approach would have been to have separated the manual FHP into a separate category. The prebarcoded and other automation mail would generally have more similar cost characteristics than the non-automation compatible mail, as Prof. Roberts agrees. Tr. 23/8404. In the meanwhile, Prof. Roberts's additional outputs results are rendered useless by his output characterizations.

V.D.2. Correctly Incorporating the Manual Mailstream in the Roberts Models Shows Expected Behavior of Workhours, though the Models are Inefficient.

As an indication of the effects of correctly distinguishing manual from automation FHP in the letter-shape models, I estimated two-output models using, instead of incoming versus outgoing FHP, manual versus automation FHP. The principal results are shown in Table 9, below. An interesting feature is the large and negative but statistically insignificant elasticities of automated letter

operations' workhours with respect to the manual FHP. Compared to the Postal Service models (USPS-T-12 at 95), Prof. Roberts's FHP-based models are very inefficient at showing that there is statistically zero elasticity of automation hours with respect to manual output. When the manual FHP are eliminated from the specification, the automation FHP elasticities are considerably better-behaved, suggesting that there is a significant efficiency cost to estimating the superfluous manual output effect in automated operations.

Table 9. Principal Results for Roberts Letter Models with Manual and Automated FHP "Outputs"

Cost Pool	Manual Variability	Automation Variability	Total Variability	Variability-Just Automation
Manual	0.32 (0.29)	0.30 (0.20)	0.61 (0.17)	n/a
Agg BCS	-0.57 (0.44)	1.51 (0.30)	0.94 (0.25)	1.18 (0.11)
OCR	-0.58 (0.68)	0.84 (0.50)	0.26 (0.40)	0.49 (0.27)

VI. Prof. Roberts's "Proportionality Assumption" Analysis Repeats Many of Dr. Neels's Errors from Docket No. R2000-1 and Fails to Demonstrate True Violations of "Proportionality" Between Piece Handlings and Delivered Volumes.

VI.A. The "Distribution Key" Method is Necessary, Appropriate, and Used by Prof. Roberts.

Prof. Roberts joins Dr. Neels in broad criticism of the "proportionality assumption" underlying the "volume-variability/distribution key" method for computing volume-variable costs by subclass. OCA-T-1 at 9-10. Like Dr. Neels in Docket No. R2000-1, Prof. Roberts advances an analysis of the relationship between MODS FHP and TPF as purportedly showing violation of the "proportionality assumption." I discussed this extensively in Docket No. R2000-1 (Docket No. R2000-1, USPS-RT-6 at 10-22, 26-28), and from a theoretical standpoint, that discussion also applies to Prof. Roberts's critique.

Ironically, Prof. Roberts in this proceeding and Dr. Neels in Docket No. R2000-1 are united in criticizing the distribution key method while proposing distribution key methods of their own. In Prof. Robert's case, his distribution key method is outlined in the response to USPS/OCA-T1-24 (Tr. 23/8335-8340) and the mathematical details of the method were confirmed on oral cross-examination (Tr. 23/8342). While Prof. Roberts considered the mathematical explication of his method a theoretical document and insists on calling FHP "plant volume," his testimony speaks for itself in terms of the output variables he used

and the resulting costs requiring distribution to subclass. Obviously, Prof. Roberts's self-deconstructing critique would only be made worse by adding further assumptions regarding how FHP relate to his unobserved idealized outputs.

As I argued in Docket No. R2000-1, the Commission should accept that the mathematics of the distribution key method constitute a legitimate and necessary approximation to the "true" or "constructed" volume-variable costs, and do not bias the volume-variable cost calculations. The Data Quality Study agreed that the general methods for computing volume-variable cost in the Cost and Revenue Analysis are sound:

> The Postal Service uses an economically sound approach grounded in activity based concepts to determine its sub-class unit volume variable costs (UVVCs) on which Postal Rates are based. The categories of data collected and analyzed are sufficiently detailed and appropriate to arrive at the sub-class UVVCs.

Data Quality Study, *Technical Report #1: Economic Analysis of Data Quality Issues* at 32.

VI.B. Prof. Roberts's Models Purporting to Show Violations of Proportionality Do No Such Thing When Fundamental Errors Are Corrected.

VI.B.1. As With the Labor Demand Models, Prof. Roberts's Choice of Instrumental Variables is Inappropriate.

Problems with the use of quarterly dummy variables as instruments extend to Prof. Roberts's analysis of the relationship between TPF (or TPH) and FHP. Prof. Roberts's TPF model specifications mirror his labor demand models

in failing to include quarterly effects as explanatory variables, and including quarterly dummies as excluded instrumental variables. Response to USPS/OCA-T1-9, Tr. 23/8303. Prof. Roberts had not explored alternative specifications (response to USPS/OCA-T1-19(d), Tr. 23/8327), and he did not conduct overidentifying restrictions tests to show that his treatment of the quarterly dummies as excluded instruments was appropriate.

As I noted above in section V.B.3, seasonal effects in MODS data identified by witness Moden include effects on both the composition and characteristics of mail volumes. These can affect both handling patterns by operation as well as the required number of sorts per piece, and so there would be no prior reason to exclude seasonal variables from piece handling models.

Consistent with the operational expectation that seasonal effects would be present in the data, the outcomes of the overidentifying restrictions tests on Prof. Roberts's piece handling models show Prof. Roberts's treatment to be inappropriate. With the quarterly dummies dropped from the list of excluded instruments, the overidentifying restrictions are not rejected. Including the quarterly dummies shows them to be jointly significant. Please see section VI.C.3, below, for the test results.

The logical conclusion is that Prof. Roberts should have treated the quarterly dummies as included explanatory variables, and not as excluded instruments. As I show in section VI.C.3, below, Prof. Roberts's error drives his most contentious results.

VI.B.2. Prof. Roberts Misspecifies the Relationship Between TPF and FHP

Like Dr. Neels before him Docket No. 2000-1,USPS-RT-6 at 18-19, Prof. Roberts's analysis is based on misspecified models of the TPF-FHP relationship. Prof. Roberts's models attempt to explain operation-level TPF using a single, shape-level FHP variable. OCA-T-1 at 12. Prof. Roberts's model is neither operationally nor econometrically appropriate.

From an operational standpoint, Prof. Roberts's models fail to reflect basic facts of mailflows and the connections between mailflows and piece handling measures. Because, in general, various components of FHP would be expected to make different contributions to particular operations' TPF, employing only shape-level aggregates of FHP is conceptually inadequate for describe operation-level TPF. Please see Appendix A for additional discussion of the mathematical relationship between TPF and FHP.

Some components of FHP will, in fact, have little or no effect on certain operations' TPF simply because the pieces do not flow to certain operations. The canonical example, which Prof. Roberts acknowledges, is that pieces recorded as manual FHP will not flow to automated operations. Response to USPS/OCA-T1-7, Tr. 23/8299. Nevertheless, Prof. Roberts's models incorrectly force manual and automated FHP to have the same effects on automated piece handlings, even though the true effect of the manual FHP is zero.

The misspecifications arise as Prof. Roberts's piece handling models are:

(1) $\ln TPF_{ij} = ai + b \ln FHPshape_i + g \cdot Xi + e$

when the true relationships are given by:

(2) TPFij= b1i FHP*ij + b2i FHP*i,-j

where FHP*ij is true FHP in operation j, and FHP*i,-j represents a vector of true FHP in operations from which subsequent handlings may flow to operation j. In the case of automated operations, it is possible to write:

(2a) TPFij= b1i FHP*ij + b2i FHP*i,-j + b3i FHP*i,man,

where FHP*i,man is (true) manual FHP. Mailflows imply a null hypothesis of b3i=0 – i.e., since manual FHP do not flow to automated operations, they should not cause TPF in automated operations.

In general, equation (1) is inappropriate because equation (2) suggests that same-operation FHP and other-operation FHP may have different effects, and equation (2a) shows that in automated operations the manual FHP would not be expected to affect TPF at all, when contrary to fact Prof. Roberts's models force manual and automated FHP to have the same effect.

In defense of his methods, Prof. Roberts suggests that his models reflect the relative contributions of the components of FHP. Response to USPS/OCA-T1-19, Tr 23/8326-8327. Prof. Roberts's claim does not rescue his approach. It is true that if the components of FHP (FHPj, FHP-j, and/or FHPman) could be expressed as constant proportions of the shape-level aggregates, then equation (1) could, in principle, estimate the joint effect of the FHP variables in equations (2) or (2a). However, Prof. Roberts's own data analysis shows shifts in the composition of FHP, notably away from manual FHP, even over the relatively short time periods under consideration. The only reliable basis for Prof. Roberts to assert that accommodating mailflows in his models is irrelevant would have

been to have estimated less-restricted models such as equation (2) and shown that the results were substantially unchanged versus his restricted specification. He did not do so.

VI.B.3. Valid Shape-Level Specifications Show No Significant Violations of "Proportionality" Between TPF and FHP.

For shape-level aggregates, it is possible to write a relatively simple equation relating TPF and FHP without the need to account for differential effects on handlings from various portions of the mailstream. Docket No. R2000-1, USPS-RT-6 at 16-17. Using either Prof. Roberts's specification from his threestep.do program in OCA-LR-2 or a specification allowing separate effects on TPF from automated and manual FHP, there is no evidence that the TPF-FHP elasticities are statistically different from unity. See Table 10, below. These results are consistent with similar non-instrumental variables results from Docket No. R2000-1. Docket No. R2000-1, USPS-RT-6 at 22.

Table 10. Results for Shape-Level TPF-FHP Models

Shape	Roberts Model[1]	Reject Over-identifying restrictions?	USPS	Reject Over-identifying restrictions?
All Letters	1.04 (0.03)	Yes	0.85 (0.13)	Yes**
All Flats	0.79* (0.01)	Yes	1.13 (0.15)	No

* Differs from 1.00 at 5% significance level or better

** Rejects at the 5% significance level, but not the 1% level

[1] Estimated using procedures from OCA-LR-2, program threestep.do.

VII. Appropriate Methods for Applying Roberts-Style Models to Cost Segment 3.1 Can and Should Make Use of Existing Data Systems and Analyses

VII.A. The Subclass Volume-Variable Cost Method Described by Prof. Roberts Clearly Applies the Distribution Key Method

While Prof. Roberts has been highly (if inappropriately) critical of the "distribution key" method for calculating volume-variable cost by subclass in the Postal Service CRA (OCA-T-1 at 11), when Prof. Roberts articulated a method for implementing his models in the mail processing cost calculations, the method he described was itself an example of the "distribution key" approach. Response to USPS/OCA-T1-24, Tr 23 8337-8340. The elasticities Prof. Roberts estimates in his empirical model do not identify volume-variable costs by subclass, so he posits distribution keys that assign the volume-variable costs associated with his models' FHP outputs to subclasses of mail. Moreover, since plants' sorting volumes are distinct from the "delivered volumes" Prof. Roberts considers to be the final outputs of the Postal Service, Prof. Roberts's method clearly involves a "proportionality assumption" of its own.

What has happened is that Prof. Roberts has stumbled upon a major difference between his theoretical framework and his empirical implementation. Prof. Roberts's theoretical framework assumes the ability to measure costs for a large number of distinct mail products (based on mailer preparation level, the depth of sortation in Postal Service operations, subclass of mail, and other relevant cost-causing characteristics), when as a practical matter the data and econometric models Prof. Roberts deploys can only measure costs for a relative

handful of the relevant categories. Prof. Roberts finds himself in the very common position of not being able to empirically implement an idealized "constructed marginal cost" model, and needing to rely on the assumptions he has criticized in order to proceed with a feasible approach.

Put another way, Prof. Roberts's explication of his model in terms of a single product (Roberts 2006 at 12-26; seminar transcript at 20-22) was not a useful simplification of reality, insofar as the mail processing costing exercise fundamentally involves computing costs for multiple products. Prof. Roberts seems to have misled himself into believing that a "constructed marginal cost" method was feasible when it is not; as the punchline of the old joke about the economist and the can of food on a desert island goes, by assuming away the Postal Service's products, Prof. Roberts assumed his can opener.

As was discussed, above, the "distribution key" method is both necessary to implement subclass-level volume-variable costs and conceptually reasonable in that given a suitable choice of the distribution key, it constitutes an approximation to the "true" volume-variable costs. Thus, the practical necessity of employing the distribution key method does not involve any material compromise of costing accuracy.

VII.B. Limitations of Prof. Roberts's "Output" Measures Suggest that IOCS-Based Distribution Keys are Appropriate to Represent Cost Differences Within Volume Categories

Prof. Roberts's theoretical model assumes the ability to measure costs for a large number of distinct mail products (based on mailer preparation level, the

depth of sortation in Postal Service operations, subclass of mail, and other relevant cost-causing characteristics), when as a practical matter the FHP data and econometric models Prof. Roberts deploys can only measure costs for a relative handful of the relevant categories. (See, e.g., OCA-T-1 at 43-44.) Therefore, not all of the relevant differences in resource usage will be reflected in the marginal costs associated with the various FHP measures in Prof. Roberts's empirical models. Using distribution keys of pieces by subclass, which assumes uniform marginal costs within each pool of volume-variable costs, will thus tend to overstate costs for low-cost mail that is either more presorted or processed to less-than-average sort depth, and conversely will underestimate costs for higher-cost mail categories.

Should the Commission choose to adopt a variation of Prof. Roberts's empirical models, it should use IOCS-based distribution keys to deal with relative workload differences within the output categories in Prof. Roberts's models. As the Commission observed in Docket No. R97-1, the IOCS-based subclass distribution keys may be regarded as measuring the distribution of pieces processed in the operation weighted by labor usage (PRC Op., Docket No. R97-1, para). Thus, the IOCS distribution keys would effectively de-average costs within Prof. Roberts's FHP categories and avoid the problem of under-distributing volume-variable cost to mail categories with relatively high resource usage.

VII.C. Using Prof. Roberts's Single Output Model is Conceptually Inappropriate and Does Not Solve Distribution Key Issues

Prof. Roberts suggested the possibility of employing his single-output models as a means of circumventing limitations in the available volume data. (Tr. 23/8440.) This approach would enshrine the least appropriate form of Prof. Roberts's models, both from an econometric and operational perspective, as discussed above. Nor are they even Prof. Roberts's recommended models, which for letter-shape operations are two-output models similar to those presented in his March 2006 paper. Response to USPS/OCA-T1-8, Tr. 23/8300. Nor does it actually solve any distribution key data availability problem. Therefore, the Commission should not adopt this approach.

Much of the point of Prof. Roberts's March 2006 two-output models was to begin to recognize resource usage (and hence cost) differences between components of shape-level FHP. Roberts 2006 at 4-6. It follows that employing variability models based on *more* aggregated FHP *increases*, not decreases, the need to deaverage the FHP marginal costs.

Beyond the need to deaverage the FHP marginal costs implied by Prof. Roberts's models, Prof. Roberts's suggestion that subclass volume measures could be used to distribute costs derived from more aggregated FHP-based models is incorrect. This appears to result from Prof. Roberts's mistaken belief that FHP and ODIS-RPW volumes differ primarly in that FHP data do not offer subclass detail. Response to USPS/OCA-T1-4.b, Tr. 8293. In fact, since the ODIS-RPW volumes include pieces that bypass sorting operations, and in some cases even mail processing plants in their entirety, employing such data for

sorting operations' distribution keys would inappropriately distribute costs to mail that bypasses sorting operations. Conceptually, IOCS is a superior source of distribution key data since it is possible to identify sets of "handling mail" tallies representing pieces processed in the modeled operations.

VII.D. Using Existing Cost Pools and Distribution Keys with the Roberts Model

The cost distribution method described by Prof. Roberts may be applied with relatively modest changes to present volume-variable cost calculations. In particular, since Prof. Roberts's labor demand equations correspond to cost pools, his method permits (and, indeed, requires) the retention of the MODS cost pools he models. The distribution keys may employ IOCS tallies corresponding to sets of operations used to aggregate the FHP. For Prof. Roberts's two-output (incoming and outgoing) model, the subclass volume-variable cost calculation Prof. Roberts proposes is:

$$RVVC_{sj} = C_s \eta_{s,IN} d_{s,IN,j} + C_s \eta_{s,OUT} d_{s,OUT,j}$$

Where s indicates the shape of mail, and j indicates subclass; Cs is the cost for the shape s operations, the terms η are the estimated elasticities and the d's are distribution key shares. Since Prof. Roberts calculates his shape-level elasticities as:

$$\eta_{s,IN} = \sum_{i \in s} \phi_i \eta_{i,IN} \text{ and } \eta_{s,OUT} = \sum_{i \in s} \phi_i \eta_{i,OUT},$$

where ϕ_i is the cost share for operation (i.e., cost pool) i associated with shape s, and the elasticities inside the summation are from Prof. Roberts's cost pool-level labor demand equations. Note that for cost pool i,

$$C_i = \phi_i C_s .$$

Combining results, Prof. Roberts's formula is equivalent to

$$RVVC_{sj} = d_{s,IN,j} \sum_{i \in s} C_i \eta_{i,IN} + d_{s,OUT,j} \sum_{i \in s} C_i \eta_{i,OUT}$$

That is, the pools of volume-variable costs to be distributed (the summation terms) are the product of the existing cost pool dollars and the elasticities from Prof. Roberts's individual equations. The result is then distributed to subclass on the appropriate distribution keys.

VIII. Conclusions

When correctly applied, the analyses presented by Prof. Roberts and Dr. Neels provide no indication that mail processing variabilities materially exceed 100 percent at the cost pool level or for larger aggregates of mail processing labor costs. These results are broadly consistent with those presented by the Postal Service for nearly ten years. Accordingly, the Commission should adopt econometric models consistent with these results. While the Postal Service models are most consistent with the structure of operations and the most demonstrably robust of the methods under discussion, I would regard Prof. Roberts's March 2006 models as updated in USPS-T-12 as an acceptable alternative.

Outside of sorting operations, variabilities should be determined using the Postal Service method of computing a weighted average of the the sorting operation variabilities, or alternatively using the results from the full-plant version of Dr. Neels's models, as presented in Section IV.B.2, above.

Should the Commission not accept an econometric variability model, it should recognize that certain costs, notably for setup and take-down activities in sorting operations, are not volume-variable. The method described in my response to MPA-ANM/USPS-T12-2 (Tr. 10/2527), is appropriate for identifying the costs to be treated as non-volume-variable. As described above, the Commission should decline to assign the non-volume-variable costs to subclass

other than via a theoretically correct treatment of inframarginal costs, such as that incorporated in the Postal Service incremental cost model.

In the event the Commission adopts no econometric variability model, the Commission should note that all of the alternatives under consideration are applications of the panel data econometric framework, and direct that subsequent analysis employ that approach as the correct econometric methodology, as recommended by Prof. Greene in Docket No. R2000-1 (Docket No. R2000-1, USPS-RT-7 at 5, Tr. 46E/22040).

Finally, should the Commission adopt econometric models derived from Prof. Roberts's and/or Dr. Neels's work, it should continue to distribute costs to subclass according to the present IOCS-based methods. This is necessary to ensure that cost differences not captured in the FHP outputs of those models are properly reflected in subclass volume-variable costs.