OFFICIAL TRANSCRIPT OF PROCEEDINGS
BEFORE THE
POSTAL RATE COMMISSION

In the Matter of:)
)
WORKSHOP BY PROFESSOR MARK)
ROBERTS ON HIS ECONOMIC)
FRAMEWORK FOR MODELING MAIL)
PROCESSING COSTS)

Date: March 14, 2006

Place: Washington, D.C.

Pages: 1 through 85

HERITAGE REPORTING CORPORATION

Official Reporters
1220 L Street, N.W., Suite 600
Washington, D.C. 20005-4018
(202) 628-4888
hrc@concentric.net

POSTAL RATE COMMISSION

In the Matter of:)
)
WORKSHOP BY PROFESSOR MARK)
ROBERTS ON HIS ECONOMIC)
FRAMEWORK FOR MODELING MAIL)
PROCESSING COSTS)

Suite 200
U.S. Postal Rate Commission
901 New York Avenue, N.W.
Washington, D.C.

Tuesday, March 14, 2006

The above-entitled matter came on for hearing pursuant to notice, at 12:35 p.m.

PRESENTATION BY:

MARK J. ROBERTS
Department of Economics
University Park, PA 16807

MODERATOR:

SHELLEY DREIFUSS
Office of the Consumer Advocate
Postal Rate Commission

APPEARANCES:

On behalf of USPS:

THOMAS BOZZO
Christensen Associates
4610 University Avenue, Suite 700
Madison, WI 53705

FRANK R. HESELTON
United States Postal Service
475 L'Enfant Plaza, S.W.
Washington, D.C. 20260
(202) 268-5204

APPEARANCES (continued):

ROBERT W. MITCHELL
(302) 340-1254

STUART ELLIOTT
for MPA, Pitney-Bowes
SLS Consulting
1920 N Street, NW.
Washington, D.C. 20036
(202) 349-7323

PETER HUME
7837 Southdown Road
Alexandria, VA 22308
(703) 768-2018

JOEL THOMAS
National Association of Presort Mailers
1195 Mace Road
Annapolis, MD 21403
(410) 990-1180

MARC SMITH
United States Postal Service

BARRY WALSH
United States Postal Service - Operations

TED PEERSALL

NALAN MONCH
Postal Rate Commission

JOANNE MONACO
Postal Rate Commission

LARRY FINFIELD
Postal Rate Commission

CHARLIE ROBINSON
Postal Rate Commission

JOHN WALLER
Postal Rate Commission

APPEARANCES (Continued):

ERIC KOETTING
United States Postal Service
475 L'Enfant Plaza, S.W.
Washington, D.C. 20260

P R O C E E D I N G S

(12:35 p.m.)

MS. DREIFUSS: This is the afternoon session with Professor Mark Roberts. This a more informal session that we're calling a workshop. We've got a lot fewer people here this afternoon than we did this morning and I think most of those who are here this afternoon are more proficient in econometric analysis and they will have questions of a more technical nature.

We're ready to start with questions. Who would like to go first?

Again, since there are people listening over the Internet, please identify yourself and your party. Thanks.

MR. BOZZO: Tom Bozzo with Christensen Associates representing USPS.

I just want to turn back to a question that Mr. Koetting asked you in the morning session about your characterization of outputs as two groups of the letters receiving the initial sort and the letters receiving the final sort. You characterized the plant output as the bundle LI and LF.

In your 2002 model, you define the output as the number of sorted pieces and I believe Mr. Koetting

asked you if you were aware that the sum of LI and LF was not necessarily equal to the unique number of sorted pieces in the plant.

Would you agree that if you considered the bundle LI, LF as you describe it on page 3 of the handout versus an index number LI plus LF that the latter is the more general characterization of the plant output, that the pair of outputs is a more general characterization of output than the sum of the WO?

MR. ROBERTS: I guess the first point to make is that output here is always -- in my view of the world, output is always sorted pieces and output never changes as a result of what the plant does, okay? It's the number of pieces of mail that arrive at the plant.

Now, they can arrive with different characteristics, different amounts of presorting, bar coding, et cetera. They can leave with different characteristics, different depth of sort, but the number of pieces of mail that comes in never changes as a result of what the Postal Service does within the plant.

So all I'm doing is taking that total number of letters and saying we can divide it up into boxes.

The sum of the boxes will always add up to the total number of pieces that come in, so nothing is ever being created or destroyed here.

Now, is it more general to represent output as a quantity or number of pieces of mail with one set of characteristics and a different quantity with a different set of characteristics?

Sure. That's more general than adding the two together and saying, no, they really all have the same characteristics.

So I think of the bundle LI, what I've called LI and LF in the model, those are two mutually exclusive categories. The sums of the pieces in each those two categories add up to the total number of pieces that are there in the plant.

MR. BOZZO: Now I want you to just consider the LI part of the bundle which is represented as a stylized fact of the data that there's relatively little secondary handling in what you'd consider the initial or outgoing sorting function. So let's assume for the sake of discussion that in the initial sort process that the Postal Service can in fact sort the mail initially to all of its destinations in one pass.

Why wouldn't it be true that MODS TPF wouldn't measure LI in that case?

MR. ROBERTS: So give me this case again? What are you saying?

MR. BOZZO: I'm saying that assume for the sake of discussion that the outgoing or initial sort can take place in one sort pass.

MR. ROBERTS: So all that the plant is doing is taking in an initial number of letters, sorting them one time and they're all sent out at the same depth of sort, the same characteristics on the outgoing --

MR. BOZZO: They are sent out at some depth of sort that can be achieved in one sort pass. Exactly what depth of sort we don't have to make assumptions about that. I'm just saying whatever set of output sorting characteristics they have, they can achieve it in one pass.

MR. ROBERTS: So each letter is handled one time?

MR. BOZZO: Right.

MR. ROBERTS: And whether you want to count it as coming in the door here and counting it when it arrives in the door or whether you want to count it when it moves one time through one machine and that's it, sure, you need the same number in that example.

MR. BOZZO: And is it your understanding

based on the MODS definition that TPF for an outgoing operation is about equal to LI as long as there's not too much secondary sorting?

MR. ROBERTS: I never really thought much about it because I've just never used -- I've never viewed TPF as that useful a measure, other than a measure -- in automated operations as a measure of machine time. I think there is something to that. I've never really thought about it as a measure in the kind of way you're describing it.

What we want is a measure of the number of pieces of mail. We want volume measures.

Now, there may be some cases where there's other ways to measure volume. I think that's great. I'm happy to think about it. I don't know how general that is.

MR. BOZZO: The other thing that I would like to ask -- Tom Bozzo still -- regarding your measuring the cost drivers section of the presentation, Section 2 where you claim that TPF is not an output measure, you note that TPF is proportional to hours of machine time using the operation.

Is it fair to say that you got that relationship out of an equation in my R2K5 testimony?

MR. ROBERTS: Yes, that's in your paper.

MR. BOZZO: And are you aware that if you add another technical parameter, which is to say the number of people that it takes to staff a machine, you can establish the proportionality of portion of work hours and TPF?

MR. ROBERTS: I don't know what you mean.

MR. BOZZO: Okay. If you take basically a throughput index and multiply it by TPF or literally if you divide TPF by throughput you get machine time. Is that correct?

MR. ROBERTS: Mm-hmm.

MR. BOZZO: Now, if you take TPF, divide by throughput, take that quantity and multiply it by the number of people that it takes to staff a machine, that gets you a measure of a portion of the labor hours of the operation. Isn't that right?

MR. ROBERTS: Okay.

MR. BOZZO: So if TPF is proportionate --

MR. ROBERTS: At a constant? Is that the thing that you're multiplying it by? Is it something that's always constant, when the machine is running?

MR. BOZZO: Well, it may --

MR. ROBERTS: It is on your TPF side, right? It's constant when the machine is running. Is the

labor relationship constant when the machine is running?

MR. WALSH: This is Barry Walsh, Postal Operations. Essentially, it is.

Most all the machines are assigned what's called a labor index and there's a standard staffing for it.

Now, occasionally it may be that they run short staff or something, but of the most part, yes, it's a constant relationship.

MR. BOZZO: Yes. It's more constant for operations like BCS where there's a single person feeding a machine, a single person working the sweep side. There are machines like the AFSM which have variable capacity based on, say, how many of the input stations are worked, where things are a bit more complicated, but taking BCS as an example -- well, in fact, even if the staffing index is non-constant, isn't it true that the staffing is still dependent on the number of sortations that the machine has to perform to complete the sort plan?

MR. ROBERTS: I would assume that the longer the machine runs the more people you need to operate it. Sure.

MR. BOZZO: I guess the question is if TPF

is proportional to machine time and work hours, basically depending on exactly which set of technical parameters --

MR. ROBERTS: Do you see that in the data? Do you see TPF proportional to work hours in an operation? I mean, that's your productivity measure, right? That has a wide range of outcomes.

MR. BOZZO: Sure. There is a distribution of productivity, but of course the elasticity of work hours with respect to TPF, which is what the Postal Service's models purport to measure, it also varies but is on average something in the vicinity of 85 percent, depending on the operation.

I guess the factor of proportionality doesn't matter. I guess my question is if TPF is proportional to some degree to both machine time and some work hours, how is TPF only a measure of capital input? That's my last question.

MR. ROBERTS: So I think the idea, the way I was viewing it, is that people are clocked in to operations, right? They're clocked in, they're producing man hours in those operations regardless of whether the machine is processing mail this instant or not. If the machine is operating -- if the machine is turned on but it's not processing anything, you're

still accumulating hours in the operation, right? And the capital services say here's exactly how much the machine is running, here's exactly how much work or input the machine is providing. So I just viewed it as you could have man hours that are operating the machine regardless of whether it's counting a piece of mail or not and then you have an independent measure, you have an additional measure now of just how many pieces the machine counted, how many minutes or how many seconds that machine was operating.

I don't see where those things are nailed down in a fixed proportion and your data shows they're not, right? Because that's just your productivity measure. So we know that those things vary.

So I viewed one as a measure telling me just how much work the machine did, the other is telling me how many man hours I've got clocked into the operation.

Sure, they're going to be related. Absolutely. If the machine is used more, you're going to see more labor hours and you're going to see more TPF. Absolutely there's going to be a strong positive correlation, but I think it's a strong positive correlation between the two inputs that are being used to sort this stack of letters that arrived at the

machine and I don't see any measure there of output where output is number of letters sorted. What did they look like coming in, what did they look like going out, there's no measurement of that at all. It's purely a measurement of how many seconds the machine is running and how many man hours are clocked into the operation. I see those as correlations between two inputs.

MR. BOZZO: If you don't see any output in piece handling, then what do you see as the purpose of operating the machines?

MR. ROBERTS: Output as we want to measure output for the purposes of measuring marginal cost. We're trying to quantify the cost effect of an expansion in output. When we say that, we mean the volume of mail and so what I'm always looking for in this data, what I'm always looking for in my model is how do we pin down the volume of mail here? And I don't see the volume of mail in the plant or anywhere else when I look at these two variables for a BCS operation. I don't see where that's related to the volume of mail that's moving through the plant.

Sure, they're giving us information about what's going on in the plant. Absolutely. In fact, I spent a lot of time saying what does TPF measure? How

can I even think about this? How do I approach it? Because, sure, it's got information in it, right? I mean, it's certainly varying with something that's going on in the plant and there's information there. How should we use it? How should we fit that into the general framework? And that's where I start thinking about it as an input, rather than an output.

What it goes back to is when I talk about output I always mean volume of mail because ultimately that's what we're trying to measure the cost of and everything else, there can be intermediate steps and your cost driver methodology takes advantage of this. Intermediate steps that are linked in a way to volume, I don't think linked as cleanly as your model assumes, but that are linked to volume, but they're still not volume and ultimately we want the marginal cost of the cost of an additional piece of mail.

MR. WALSH: Barry Walsh, Postal Operations. I'd like to -- maybe I can shed some light on what's going on and what the link is between TPF and volume as you refer to it.

One of the main things that's going on that causes this wide variation in productivity that you referred to in looking at the data, say, in the DBS, to take a simple case, is the impact of having to

change the scheme on the machine or the sort plan.

If you're doing incoming mail, your final final, you have a lot of scheme changes because you're sorting to many very separate places, you sort the mail to Reston, you sort the mail to Alexandria, you sort the mail to Arlington, and you have to put on a new scheme for each one of these. It doesn't make any difference how much mail you've got, as long as you have any mail for Alexandria, you have to run the Alexandria scheme.

The number of the scheme changes and the time that you take for each of these scheme changes accounts for that difference in predication that you're seeing.

Once you account for that, you have a fairly direct relationship between TPF and total volume because you simply have to look at the number of passes it takes, adjust for the scheme changes and you've got your total volume and that's the way it actually works in practice. When we do budgets or we do planning, that's the kind of thing that we plan.

MR. ROBERTS: I understand that and that makes a lot of sense. Once the scheme is set -- maybe this is what you're saying -- once the scheme is set and you know how many passes each letter has to take

to be sorted, there's a direct link between TPF and volume.

Now, the problem is what are we using? We're using quarterly data, we're using data that's aggregated up across many, many, many hundreds or thousands of scheme changes, right?

So what we're seeing at the quarterly level is what comes out as the relationship between hours and TPF.

So I'm not surprised that productivity varies. When I look at the productivity measure, actually that makes a lot of sense, this being one reason, but what Tom was trying to push is the idea that there ought to be some fixed relationship between these two things and I'm just saying it ain't in the data, when you look at quarterly data.

Now, maybe if you go down to the level of we're running one scheme this evening and it's not going to change, then maybe there is a fixed relationship, but that gets wiped out through the aggregation that we go on, so we're only going to see that at a more aggregate level in the kind of data that we're using.

MR. WALSH: I think you can actually see this when you look at the difference in productivities

between what you refer to as initial and final, we call them outgoing and incoming sorts. On your outgoing, you do have these long, long runs. You have a scheme that you're running for the whole country and you run it for a very long time. On the incoming, you have all these multitude of scheme changes. So we tend to end up with a lot more economies of scale on the incoming than we do on the outgoing.

MR. ROBERTS: That's a good argument for separating the output, not saying that there's just plain vanilla FHP that goes through the plant, but that there's FHP of different types going through the plant. I agree with that. That point was brought up in my last seminar and I was aware of that at the time and I actually responded that this -- kind of breaking them down into these different categories is a useful way to go and so what I tried to do this time is say can I make some progress on doing that with the MODS data? I think, yes, I've made some, but I think it's a reasonable starting point and that's exactly what the disaggregation into these two types of output is trying to account for. And I would expect to see different labor implications for the outgoing processing than for the incoming processing because they're different animals, they're doing different

things. Yes, Intervention motion happy with that.

MR. HUME: I'm Peter Hume. I'm a private consultant. I've worked with the Postal Service for a long time.

In light of those differences in productivity, I can recall when I first started working on postal stuff there was virtually no automation and most sorting was done manually on cases and so the direct relationship between piece volume and man hours was pretty self-evident. The differences in productivity and in marginal cost among facilities was extremely large, a factor of three-to-one in unit costs sometimes. You could actually see that. And the point is that the data were not MODS data in those days. The data were really actual pieces and man hours because that's what we used to count.

Further to that point, I wonder if you have sufficient data that you can disaggregate your MODS data among facilities and in particular facilities which have unusual characteristics or differential characteristics. There are some facilities which, for instance, like in Washington work near a place where there's a lot of paperwork comes out and there are other places where there's a lot of bulk mail comes

out and these disaggregations might give you some insight as to whether you have enough variables in your model.

MR. ROBERTS: I don't know anything about these plants other than -- including their geographic location -- other than what's reported in the MODS data. That's been one of the frustrations I've found in working with the MODS data. I'd like to know more about particularly geographic locations on some of these plants, but that's not something that's available to me.

Now, it is available to the Postal Service. I guess it could be used, but it's not something I have access to.

MR. HUME: I had tried to do that in the past, not having full access to the data, but very frequently I found that once you start disaggregating, your models fall apart because you don't have enough degrees of freedom to handle the models you have and so practically it's a difficulty, but I would certainly advise if you can get disaggregated data that would be very helpful.

MR. ROBERTS: Well, this data is at the level of the processing plant in a quarter. That's the unit of observation, so 350 approximately plants,

aggregated over 12 weeks of data. So it is disaggregated in that sense.

Your question was one of are there things that are atypical that are going on in some of these plants. Yes, there certainly are. I pointed to one in the seminar where there are some plants, there's about 40 to 50 plants, that have no automated flat sorting that I can see. When I look at the automated flat sorting operations, there are no hours reported, there's no TPF reported, so I assume it's not there. I'm joust looking at the data and saying, okay, I don't see any report, I assume those operations don't exist. I'd really like to know that for certain.

MR. HUME: But my point is that if you do know that, then you can do that to gain confidence in your models. They should be representing that correctly and if they're not then you've really got a problem.

MR. ROBERTS: I completely agree. One of my recommendations at the end was that I think we really ought to move to standardizing the set of plants we're looking at for just this reason. There are plants that are clearly not doing the same set of operations as others. I don't understand why, but certainly the

postal staff does, and I think it would be useful to sit down and go through the plants one by one, 350, you've only got to do it once, go through them one by one and say is this a plant that is a representative plant in terms of the mix of technologies that it's using, in terms of the volume of mail that it's handling, in terms of the service area, the geographic area that it's serving. Is this somehow a representative plant that we think of as being one of the plants that's carrying the workload for sorting the mail?

If we came up with 200 plants that looked like that or 250 plants, I would be much happier using a data set like that than I would be a data set with 350 plants where those last 100 or 150 are quite odd. I don't feel comfortable using that.

What I have at my disposal to look at is only the reported operations, but I think it would be nice to standardize the set of plants. I think that would help.

MS. DREIFUSS: Would you identify yourself, please? We know you in the room, but on the Internet --

MR. PEERSALL: My name is Ted Peersall again.

I've looked at the MODS data that we've had from time to time here and one of the things I've noticed is that from quarter to quarter part of the MODS data is the number of zip code delivery offices that are being served by each plant and one of the interesting things about this is that it changes quite a bit from quarter to quarter.

What that suggests to me is that the mail can be moved among plants, it can be processed, the mail destined for a particular zip code can be processed at different locations at the option of the Postal Service.

If that's the case, then a model such as yours that takes the workload at the plants essentially as exogenous isn't quite right.

Have you given that any thought, looked at any of this?

MR. ROBERTS: Well, I guess the way I would approach it, Ted, is saying if a plant was receiving mail from an area that it didn't normally sort for, I don't know, to use an example, there's a problem in Richmond and some of the mail is shifted up to Northern Virginia to be sorted instead, then what's relevant in our model and in the Postal Service model, too -- well, I'll speak for my model. What's relevant

in my model is that we see an increase in volume of mail in the plant and it really doesn't matter whether that mail was coming from a Richmond plant which is not doing it and shipping it up here or if it just came from an increase in collection mail in this area.

In my model, that wouldn't matter. What matters is just the number of letters coming into the plant and I'm viewing the plant -- I'm trying to model the inside of this plant, but what's coming in doesn't really matter to me.

Now, where it would matter, where it should matter, is the kind of thing I was trying to get at with the multiple outputs.

Now, if the mail that was coming up from Richmond was already sorted to a different level than the collection mail that was coming in, I would want to be able to distinguish that. I would want to put those FHP counts in different boxes but it would still be ultimately -- at the bottom line, it would still just be a count of pieces that come in and that's what I would want.

I'm not sure how I would use or how I would even think about mail coming from Richmond being different from mail being collected in this area or if that would really be important.

MR. PEERSALL: Let me point out that if you use the elasticity estimates as system variabilities, then you run into a problem. If you treat them as just plant level variabilities, then there's no problem.

MR. ROBERTS: So you're talking about a process where you aggregate over plants to try and count what's going on on a system wide level. Yes. Yes. I could see where that's going to be an issue there. That's not what I'm doing.

MR. BOZZO: Tom Bozzo again. But isn't it fair to say that you really don't treat your output measure as exogenous because of your use of the instrumental variables procedures? That is, whether the endogeneity is coming from the Postal Service's choice of how to allocate mail across facilities versus measurement error, the instrumental variables problem basically addresses both endogeneity from both sources? Is that approximately true?

MR. ROBERTS: Well, depending on what you use as your instruments. I mean, different sources of endogeneity lead you to different instruments. And I talked about that quite a bit in my old paper. I haven't revisited that issue here. There is an example in my old paper about, you know, gee, if

indogeneity arises from differences across plants, you need a different kind of instrument than if it arises from instrument error. So, no, I wouldn't say that. There's no such thing as vanilla endogeneity. You've got to know what's the reason for endogeneity.

Now, it's true endogeneity has the same implications for the coefficients, but how you go about correcting it depends on what the source is.

That's why I think the instrument that I used -- basically, I'm just using the FHP count of the other shape -- is one I think that works well and satisfies the requirements well for the measurement error problem. I would have to rethink it if the endogeneity came from a different source, I would rethink using that instrument. I think it's a good one for the measurement error problem, which I think is the main issue that's coming up with dealing with FHP.

Now, one aside, too. Actually, one of the things I thought about, the measurement error problem in FHP arises for a couple of reasons. One is because it's weighed. Okay. That's not exactly piece counts, it's weight, but the other is because the conversion factor is what's changing over time.

One way to dampen down some of the

measurement error problem in FHP is to update the conversion factors a lot more frequently. We saw when I used the data the first time, the '94 to 2000, I think was the last data set we looked at, the '94 to 2000, the conversion factors were changed kind of in '97, maybe, it was roughly in the middle, '98, so somewhere in the middle of the data. And you saw a big drop and I have a table in my first paper that gives you the drop in FHP counts when the conversion factors were changed.

Well, what that was saying was that all along over time this error due to incorrect conversion factors was kind of getting bigger and bigger and bigger and so a way to dampen that down is just do the sampling and update the conversion factors more often.

My impression is that's not a real complicated thing. You're doing a sample, you're taking a sample of mail, right? And you're saying how does weight convert into pieces. It's something we ought to be able to do more frequently than once every X years.

MR. BOZZO: Well, I haven't done it. It's a surprisingly large problem because you have to -- of course, you have to measure mail of all the source types on a relatively systematic basis, but I don't

disagree with that, but the economists don't always get to choose their data.

MR. PEERSALL: This is about the estimator you used. If I remember correctly from your paper, you essentially used a method that's equivalent to two-stage least squares. Is that right?

MR. ROBERTS: Yes. Two-stage least squares is an instrumental variables estimator. Yes.

MR. PEERSALL: That's right.

MR. ROBERTS: It carries a little more baggage with it.

MR. PEERSALL: The Postal Service's work used another similar estimator, K-class estimator, the limited information maximum likelihood estimator, and they are slightly different estimators and I'm sure you could have used the limited information if you'd wanted to.

Would you sort of explain the difference and defend the choice that you made?

MR. ROBERTS: Well, any time you use a maximum likelihood estimator you're making an assumption about the distribution of the error term as well, that it's normal or whatever. You don't have to do that with an IV estimator. So I think there's a little bit less in the form of assumptions.

I was trying to use the IV estimator just because I think it's just real clean in the following sense. I've got one endogenous variable, I've got one instrument, and so you kind of know what variation in the data you're exploiting. I know that I'm regressing that endogenous variable on that instrument and I'm taking the fitted value of that as my cleaned up variable. And so I can look at that and I can kind of dissect it and convince myself that there's a strong correlation between the endogenous variable and the instrument and I can look at the fitted values, I can say here's how the variable changes as a result of doing this process, here's how it changes when I put it back into the model.

So in that sense, the steps are more transparent to me and I can kind of check them as I go along and say this is reasonable, I can see why this is happening.

Now, that's not the only way to do it, but what I was trying to illustrate in the first paper was kind of the importance of the problem. I was trying to say the measurement error problem is an issue here, I realize it, here's a way of dealing with it that I think is trackable, with the MODS data it's trackable, and I was trying to illustrate in the simplest

possible way so that it wasn't getting muddied up by a lot of other assumptions and that's kind of the approach I would take. I still take.

When I view a data set like this, I always try and use simple models and simple techniques where possible because I can kind of follow what's happening.

I'm not opposed to maximum likelihood. I just thought in this case this was kind of a simpler way to go. Cleaner.

MR. SMITH: Hi. I'm Marc Smith with the Postal Service. Just a couple of comments and a question.

I think you make a lot of good points in your work and you also point out that there's a whole lot of nitty gritty out there that's hard to capture. I'll point out one area that you may want to consider, which again is -- I don't know how feasible it is to do, but remote bar coding has been now used for letters and now for flats in recent years and so some of the labor used in processing letters and flats is essentially off site and not only that but there's been a technological advance in terms of being able to use computers and so over the years computers have been able to resolve more images and there's been less

need for keyers off site, so you've got both things going on and it's -- I'll just say that both of those would present challenges to you.

The other thing I just want to mention is that there is a -- how would standardization efforts relate to this? I mean, the Postal Service tries to standardize its work so as to be able to understand the plants to be able to work with each other. I mean, each plant is constrained in a lot of ways. It really can't go on its own. Each plant works with a set of equipment and they're prescribed to operate them certain ways and so they're not so much like separate actors trying to maximize without constraints. I guess I just want to raise that as an issue, that the plants are being asked to operate in a certain way that it might be that real wages perhaps differ place to place but I'm just suggesting that there's this other issue that could be important.

MR. ROBERTS: Let me respond to a couple of those points, Marc. Thank you.

One was the remote bar coding. I saw that last time, it was there, that's not one of the operations that we've included. I think it should be. I think the remote site labor is a labor input that's used in the sorting operations and I think it should

be included as another labor input. I don't have any problem with doing that. I don't have the data to do it. Rethinking how would you include that in the model in a practical way, I haven't thought enough about it, but I think it could be done. I think conceptually I'd like to count that labor. Yes

The second one, your comment about the improvement in picking up the images, absolutely. It's really quality improvement in capital, it's probably quality improvement in software but we kind of treat software as capital that depreciates kind of fast, I guess. Ideally, we'd like to measure expenditures on capital equipment and software to operate it as part of the capital account.

Now, whether we pick up quality change in software, that's an age old question that a lot of people have dealt with. The answer is we don't do that very well in most cases. Maybe here you could, maybe because the software is specialized enough that in measuring capital you could try and control for quality improvements in software as well. But the way I would bring that into the production model that we use here is through the capital stock variables, through measurement of capital in the plant.

MR. BOZZO: It can be done. There's new

complications related to the use of remote encoding for non-letter pieces. In R97, Mike Bradley's testimony presented results that were based on regressions of remote encoding site work hours, the number of images that the remote encoding sites represented which are collected out of the computers that attempt to read the mail and send them on. So those data do exist and the analysis at the time showed there to be roughly 100 percent variability of remote encoding work hours to the number of images that were processed, but that analysis hasn't been updated in a while.

MR. WALSH: Barry Walsh. Just sort of a general question.

I understand that these estimates that you do of elasticity are estimates over time and that they're an average elasticity over four or five years, something like that. And yet the Postal Service is anything but stable in a period of four or five years. We have new equipment coming in, equipment going out. You noted that the MPBCS numbers were kind of flakey. Well, you know, that's a piece of equipment that's going out. Any time we have equipment going in and out, the numbers are flakey on it.

We also have major changes in the Postal

Service. You mentioned the AFSM coming in, that's made a big change in a lot of things. In the flats area, we have the new FSS machines introducing delivery point sequencing into plants, which will be another big change. Then we have operational changes in which they make policy changes. For example, I mentioned earlier about taking all the letter cases off the floor, getting them down to a minimum. That was partly to make room, but it was also partly as a management device to make sure that they got rid of the people that might otherwise be staffing those cases.

So given all of these changes that you have in a period of four or five years, I'm wondering how well your numbers hold up, how well you can estimate an elasticity given that the underlying technology is so non-constant and also how relevant that is to the problem of estimating an elasticity -- I guess we're trying to estimate it in the test year and yet you're giving an estimate back all these years over a period of substantial change.

MR. ROBERTS: That's a good point. The general way you approach it is by trying to control for these others things that are going on the plant through a mix of observable variables. Very important

ones in the work I do are all these capital variables. Those are things that are really changing in very systematic ways over time for a plant and controlling for those things is real important as a result.

I included these technology variables and year dummies and things like this to try and allow for these changes that are going on.

Now, does it pick them all up? No. Of course not. But the next thing you can do, and I've actually done some of this although I didn't report it in the paper because it was just getting overwhelming as it is, one of the ways I've tried to check for sensitivity of my results is to take the simple model that I've got which treats these elasticities as constants and start to introduce a small set of interaction terms, so to start to let the elasticities vary with observable things in the data. This is certainly the approach Tom has taken, much more in that direction than I've gone or that I'm really willing to go. But one of the things I did was I started to just look at the interaction between the output elasticities and the year dummies, so just a simple idea that maybe this output elasticity is changing systematically year to year as the configuration in the plant is changing.

You do find some changes in those elasticities. They weren't huge, so I didn't report them because they weren't things that were jumping out at me as really making much difference. But if I reported an elasticity of .85, maybe what you would observe -- if you estimated that separately for each year by including year interactions, you might see it go from .8 to .87 or something, move over time.

So, yes, there is some movement in these things. I think that's something that should be checked before the results are used in setting rates. It's just a way of saying is the model robust in these kind of small changes and specifications.

That said, I'm all in favor of kind of limited interaction terms that can test sensitivity of results. I think that's just a good way to check your results. I don't want to go the route of putting in lots and lots of interaction terms because I think there's a tradeoff in the other direction, that you just end up with lots of insignificant coefficients. With coefficients that can change a lot with small changes in sample, things are insignificant, and then you aggregate them all up, you get this point estimate for an elasticity and it's hard to know where that's coming from. And so I'm leery of going too far in the

direction of interaction terms, but I think it is good practice to check sensitivity of the results to an assumption like the coefficient is fixed over the whole time period. So I'm certainly willing to do some of that and I've done some of it. Mostly, it's just a matter of time, how much you can do, how much of it I can do, so I do the things -- I kind of prioritize, here's what I think the most important specification issue and I go down the list and run out of time and there's things still on the list, but we all face that.

MS. DREIFUSS: Questions seem to be slowing down. We were scheduled to go to two and, of course, we're willing to take questions until two.

Does anyone have any more questions at this point?

I do see one.

Would you identify yourself?

MR. MONCH: My name is Nalan Monch. I'm with the commission's advisory staff.

You just answered a question about how do you deal with changes over time, different characteristics among plants you deal with fixed effect variable, and one of the things that we've been seeing from other cases filed recently is how much the

Postal Service has tried to reorganize its network starting now, I guess, in 2001 it had an area mail processing initiative where they tried to consolidate the functions at certain plants, taking away, for example, outgoing sorts from smaller plants, consolidating at larger plants. Now, they're trying to reconfigure the network to apparently more closely resemble a hub and spoke configuration than what they have now. Apparently, tease are quite extensive reconfigurations that they have been doing and contemplate doing.

My question is does that make the particular role that a particular plant plays in the network so volatile that a fixed effect approach may not be valid?

MR. ROBERTS: A fixed effect is correcting for a number of things in the model. Let me back up and explain. Here's what I view the fixed effects as doing, okay? In these models. Because I use them as does the Postal Service, so I think they're appropriate to use and here's the reason, is that there are certain things about plants that make them different, that one plant, even if we took all the observable characteristics that we could, the capital stocks in particular, and we took the exact same

capital stocks from one plant and we stuffed them into another plant, would that second plant replicate what goes on in the first one?

I think the answer is probably no, it wouldn't, that there are going to be unique things about that second plant that make it different from the first one, even when we control as much as possible for the observable things that are different.

Another way of asking the question, sort of looking at the question, would be suppose we had a small plant and we had a large plant. Do we want to use the size difference in these two plants to estimate our output elasticity? Do we really want to use the fact that one plant is small, has small FHP, small hours, another plant is large, and look at the difference between those two and say, oh, yes, that's telling us about the output elasticity that we want to measure?

Effectively what we're saying is if that little plant grew up, it would look like the big plant and I think that's probably not true in most case, that when you take the small plant and you try to make it handle the mail volumes and do things the way the large plant did, it's still going to come out with a different mix of hours and FHP. And so the idea is

that the cross plant differences are not really picking up the right kind of variation in the data.

They're picking up variation that is reflecting things that are permanent differences across plants. Someone mentioned earlier in the day whether they're two-story or one-story plants. That's the sort of thing a fixed effect would control for nicely.

So what we're saying is we don't want to use that variation in the data to estimate the output elasticity. It's not the right kind of experiment in the data to estimate the output elasticity.

What we really want to estimate the output elasticity is if the plant got more FHP coming into it, more volume, what's the range of responses that that plant could make in terms of its use of hours?

So I think it's much more the time series variation in the data that we want to use for estimating the output elasticity than it is the cross plant differences.

Now, that said, both sources of variation, time variation and cross plant variation, have got useful information in them and they have some less than useful information in them and it's a matter of degree how much of one we're throwing away when we get

rid of the other.

I think a reasonable compromise is to include the fixed effects because they deal with things that are likely to be non-reproducible or non-replicable differences across plants. So that would be my argument for using them.

Now, I don't know if I answered your question or not.

MR. MONCH: I guess the thing I was focusing on is if the essential differences between plants don't seem actually to be fixed, then I guess what your response was that you sort of have an intuitive belief that the essential differences somehow are fixed even if you're doing radical reconfiguring.

MR. ROBERTS: Well, to the extent you're doing radical reconfiguring, too, it should show up in the time varying data and that's really what we're relying on to estimate these output elasticities. Think of the variation in the data, some of it's systematic and permanent across plants and some of it is time varying for both plants. If the system is under reconfiguration and volumes are being shifted from one plant to another over time, that kind of stuff is picked up in the time dimension of the data and that's what we are using to estimate the output

elasticities.

So it's really a matter of -- I guess it's a broader issue that I've wrestled with in using this data and it comes out when I talk about quarterly variation in this paper as what's the right experiment in the data, what's the right source of variation to use in estimating the output elasticity that we're after?

Ideally, the experiment we would like to do is take a plant and control the amount of mail that's going into the plant over time. So one day we get a million pieces, the next day we give it two, we give it three and we watch how the plant responds in terms of its hours used. If we could run a controlled experiment to measure the output elasticity, I think that's what we would do. We would just vary the volumes going into the plant and watch how the plant responds with hours.

So what we want when we approach a data set like the MODS data set, I approach it saying where is that kind of variation showing up in the data? Is it showing up in differences between a small plant and a large plant? No, I don't think so. I don't think that's the kind of data variation in want to use.

Is it showing up in the time series

variation for an individual plant? Yes, I think it is because now what we're seeing is, yes, a plant is in operation in a low quarter and then it moves to a busy quarter and volumes increase by 25 percent but that's reality, the plant is getting 25 percent more volume and it's dealing with it. So I look at the data, the quarterly variation, I say that's a good source of variation to use because that really is approximating the kind of experiment that we'd like to run for measuring the output elasticity, whereas I don't think the cross plant differences is the right kind of experiment.

I realize this maybe is getting a little too philosophical here, but it's just saying we want to think about where is the variation in the data and how well does it approximate the experiment we would run to estimate this parameter if we could run that experiment.

Unfortunately, as economists, we can't run the experiment, so we've got to kind of figure out where that experiment exists in our data.

I turned into professor mode there. Sorry.

MR. MONCH: I have another question if no one else has one.

MS. DREIFUSS: Could you identify yourself,

please?

MS. MONACO: Joanne Monaco and I'm with the PRC.

You know, if I remember right, looking at your results, the Houseman statistic was less than the critical value which would indicate that a random effect model would be just as applicable and I was just wondering if you tried to use that and then just compare the two.

MR. ROBERTS: No, the P values always -- we always reject.

MS. MONACO: So the null for no correlation between --

MR. ROBERTS: Yes. We're rejecting that.

MS. MONACO: You're rejecting that? Okay.

MR. ROBERTS: We would reject the random effect.

MS. MONACO: All right. And then the fixed -- okay. The fixed effect is the one to use. All right. Thanks.

MR. MONCH: Nalan Monch again with the advisory staff of the commission.

I believe I recall somewhere in your paper, not exactly where, you said it would be a legitimate thing to aggregate up the models of the various --

your shape based models by plant could be aggregated across plants in a meaningful way.

MR. ROBERTS: No, I hope I didn't say that. I don't think I said that. The kind of aggregation I was talking about was really two. One is across processing stages for a given kind of output, so we think of what I called FHP in the outgoing sort, that's a kind of output. So we can aggregate up across OCR, DBCS, manual operations and get an overall elasticity for letter sorting, so how would total labor use in letter sorting respond to an increase in FHP out? So that's one kind of aggregation that's building up to the level of letters.

The second kind of aggregation is across the outputs and across the outputs it's a different kind of question that you're asking. Now you're asking the question suppose that the total volume of mail increased in the plant. That means FHP out goes up by 1 percent and FHP in goes up by 1 percent. So all the outputs are rising, the whole vector of outputs is rising. Then, what would be total labor response? And that you can get as well. That's a different way of aggregating the elasticities.

Each of those things asks slightly different question but each of them are just recombinations of

the basic building blocks, the elasticity of each operation with respect to each output, which is what's being estimated. And I think those other things are actually -- I find them a little bit easier to think about than I do the ones for the individual sorting operations.

MR. MONCH: Let me ask a related question. You criticized the use of TPH as a poor proxy for the volume at the plant because it's not really proportional to the volume that's coming into the plant and that's what you really have as your goal, is to measure that.

Given that the objective of rate making is to apply a rate to the pieces that actually are entered into the system rather than just entered into a particular plant, is there a problem in that you have pieces that are handled by more than one plant, they can show up at one plant as an outgoing and another plant as an incoming and therefore some pieces get touched by one plant, some might get touched by three, and therefore you have a disconnect between the volume going into the system and the volume that you're actually modeling?

MR. ROBERTS: Yes, that's a great point. What's relevant for the cost in the plant, the labor

hours in the plant, is the volume that's being treated in the plant and so that's the way I'm viewing it.

Now, what's interesting is if we followed a letter through the system we could think of a letter as basically collecting costs at different stages in its lifetime, so it enters one plant and it gets an outgoing sort and it collects some cost from that. Then it goes to another plant and gets an incoming sort and it collects a different cost at that plant.

What we're estimating here, the way I approached it was I'm estimating essentially the marginal cost of an outgoing sort and the marginal cost of an incoming sort.

Now, that doesn't mean one by itself is the marginal cost of the letter. A letter is now the sum of those two things. It's an outgoing sort plus an incoming sort, so you could take margin cost and you would start to -- we're going to estimate marginal cost for different types of sorting, but then you could map those into which of these stages of sorting does a letter go through and if it's just as simple as it goes through one outgoing and one incoming, boom, then you add those up and you've got the marginal cost of a letter through that type of sorting.

I think this is actually -- I didn't have a

chance to talk about it today, but I think it's one of the advantages of this kind of disaggregation of output that I did in this case and if it could go further I think it would be nicer, but it also gives you a basis, then, for thinking about cost discounts for letters that only go through part of the sorting process, so something comes in presorted and bar coded by a mailer and completely skips the outgoing sort, then I've estimated the marginal cost of the outgoing piece of the pie, that isn't relevant for this letter, but the incoming sort costs still are. So if you could measure the marginal costs of those different pieces, you could add them up in different ways to get a cost of a presorted letter versus the cost of an unsorted letter.

I don't know if I'm making myself clear. It's something I've been thinking about. It's not made it yet into writing, so maybe my explanation is not as clear as I would like it, but I think the processing -- the way I'm thinking of the processing plant as handling different categories of mail, ultimately we're going to estimate a marginal cost of handling each of those categories. We can then describe different types of letters as different mixes of tease categories and so that will give us a way to

estimate marginal costs for different rate classes, different types of mail, discounts for presorting, this kind of stuff.

So that's where I think this could go. I've not worked out the details and I'm sure there are many, but thinking down the road.

MR. FINFIELD: Larry Finfield from the PRC. Just really an information question. I agree that the marginal cost information is key here, especially in a multi-out function where the elasticity is, as you talk about it, to a great extent a function of a share of that output and so margin costs, marginal time, is an important thing to look at and you have a way of doing that but did you sort of just in your back room ever do any calculations that just sort of --

MR. ROBERTS: Of marginal costs?

MR. FINFIELD: Yes. Marginal costs for these different operations in the incoming and outgoing.

MR. ROBERTS: I was just pressed for time to get it done and I didn't do enough. I did some back of the envelope calculations, but the thing I was running into was that the cost pools that I was working with are such a small share of cost that I was leery that --

MR. FINFIELD: For the incoming or the outgoing that you're talking about?

MR. ROBERTS: The formula that's in the paper is marginal cost is equal to average cost times these elasticities. And when I looked at average cost, it seemed like when I aggregated across by plants I had such a small share of processing costs that was actually in my data, that was actually in my pool, that I felt like I was missing something and so I wasn't willing to write them down, here's what they are. But there are costs -- you can measure these cost differences across different categories of output, but I'd have to do a lot more work to understand the size of the cost pools and the things that I'm missing when I use these cost pools, what other costs am I missing, because that seems to be the lion's share. I was uneasy about saying marginal cost is 2.16 cents or something.

MR. FINFIELD: Okay.

MR. ROBERTS: I didn't even write the number down, but it can be done. It can be done.

MR. MONCH: Nalan Monch again with the commission's advisory staff.

In your morning lecture, the point was made that the manager's discretion to direct mail from an

automated operation to a manual one was heavily constrained by a policy of removing the gauges that are used to do manual sorts and it seems that the manual sorting operation has been one of your biggest problems in modeling, particularly for flats.

I'm wondering if it would be possible to treat the manual sorting cage as though it were a technology and you could track the number of cages and the trend, what's happened to those, and perhaps add that to your model and improve the modeling of the manual operations as a result.

And even more than that, it seems from talking to managers over the years at plants that I've visited something that's mentioned quite often is space constraints and a lot of the different productivity numbers that you would compare from one plant to another seem to be related to whether they are mildly or severely constrained in terms of space. I'm not very familiar with how you've handled that with your capital variable, but I'm wondering if there isn't a more effective way to measure or compare plants in terms of how constrained they are in terms simply of square feet to do what they try to do.

MR. ROBERTS: Several points. Let me give a quick answer to a couple of them before I forget what

they were.

The one about modeling manual operations. Basically, manual letters has always, I think, been a very robust set of estimates. I haven't seen a lot of sensitivity there or as much sensitivity there as I do in the automated operations, particularly the ones that are being phased in and phased out. That's where a lot of the sensitivity comes from.

What I do see in manual flats now that I did not see using the data from '94 to 2000 is I see less output. The output variability of manual flats is much smaller than it was in the estimates from the earlier time period and I've kind of wrestled with why, you know, what's the source of that.

The one thing I can point to, and this is only part of an answer, is that a lot of the cyclical sensitivity, the fluctuations in FHP over time, and they are substantial from season to season, a lot of that seems to have been transferred from manual hours to AFSM hours and so when I look at this time period now, when I go back and look in the '90s data, I see that manual flats are pretty sensitive to the fluctuations in FHP and it looks like as FHP goes up and down quarter to quarter the manual hours are going up and down quarter to quarter.

Now, when I look at the operations in 2002, 2003, 2004, where I see that cyclical sensitivity is in the AFSM operation and I see -- just kind of eyeballing it, I see the manual hours appear to be much less cyclically sensitive than they used to be and so that's the mechanical reason that I'm getting lower elasticities now.

Now, what's that reflecting in the underlying process? Is it reflecting a shift from manual sorting in flats to AFSM sorting as the primary means of responding to volume changes? You could tell me better than I can tell you. That's what I speculate that it is.

So I think that there are things going on in the data, particularly in the flat sorting, which are responsible for these different sensitivities, different elasticities now than what I saw earlier, but I would want to explore that more before I committed to an answer.

Okay. So I got two of your questions. I forgot what the last one was.

MR. MONCH: It was about modeling technique, for example, whether you might fashion a variable that said, well, the technique is a piece of equipment and if you track what's going on with the availability of

that equipment, you might find that your model improves for manual.

MR. ROBERTS: True. I think the way I would approach that is say there's yet one more kind of capital that I want to keep track of that I haven't and that's capital in manual sorting operations and maybe the right measure to use is the number of cases or something like that.

Yes, in general, I'm all in favor of kind of disaggregating capital and including as many different types as possible, realizing there is a cost. The cost is always going to be that you're trying to ask more and more things of the same set of data and probably what's going to happen is that the precision of the estimates will start to fall off. You're going to see bigger standard errors. Maybe, maybe not. It's an empirical question, but I think the general approach of trying to control for the kind of capital that's in place in the plant is important and I agree with that and I think the more that can be done to push on that dimension the better.

MR. MONCH: I also had the question of whether your model captures constraints-based.

MR. ROBERTS: Oh, yes. No. There's really nothing there that would do that. If the total space

of the plant is fixed over time, if the total square footage of the plant is fixed, that would be taken out with the fixed effect. That's one thing that would be removed. So it would be saying, yes, I'm not using the difference between a small plant and a big plant in terms of square footage, now, I'm not using that difference to estimate the elasticity because of the way I've controlled for the plant effects.

Maybe the more relevant way to think about it is it's really something that varies over time. Some periods, it's a constraint. Some periods, it's not a constraint. The way to handle that, to try and deal with that in these kind of models, is to come up with a variable to measure. You measure what's the constraint that the manager faces. And really, you know, the way we're thinking of all these variables on the right-hand side, the way I think of them, is that they're all constraints that the manager faces.

The manager is trying to choose hours, and I think machine time, to sort mail and they face constraints on the quantity of capital that they've got to work with and they face constraints on the amount of letters that are coming in that they have to deal with and I don't have any problem adding in additional things that reflect constraints on the

managers.

Now, the ability to measure these effects, though, on quarterly data might be pretty meager. If it's something -- suppose the plant is really constrained, the cases are full five days in a quarter and so the manager does something different on those five days because of this constraint that he faces. Is that going to show up in quarterly data? Are we really going to be able to pick that out? Probably not. Probably not.

Even if we came up with a nice variable that measured, gee, for rive days of the quarter they were constrained and we put that in, would that show much action? I'm skeptical.

I think the quarterly aggregation just sort of washes out a lot of these day-to-day effects. And they're things that -- I'm not denying they're important. I bet on a day-to-day basis the managers do have to respond to these kind of constraints and things that pop up and they have to do things differently. So I'm sure they're there. It's whether or not we can measure them in quarterly aggregate data and I'm skeptical of that.

MR. MONCH: I'm reluctant to waste the time. I'll just keep asking unless somebody else has a

question.

Sometimes managers don't just have a space constraint for a few days, a few peak days of a quarter, but sometimes they simply can't install a machine that they wanted for years in a row because they don't have square feet to put the machine there and so I'm thinking that the space constraint issue is not something that's simply short-run.

MR. ROBERTS: There, the way we would approach it is we would see a different configuration of capital in this plant than we otherwise would see if they had more space. We would see an extra AFSM machine that we're not going to see because they don't have the space for it. So we're going to see different capital. We can measure that. We're going to see different hours in the sorting operations. We can measure that. So were going to be able to observe in the quarterly data that this plant does look different than an otherwise identical plant, a plant with the same volume of mail but more square footage and therefore another piece of capital equipment.

I think those are the kind of differences we can get a handle on using the MODS data, using the quarter data, because they're going to show up in the amount of capital that's there and the hours and the

different sorting operations. The things I think that are harder to pick up are these day-to-day shocks that might be a big problem at some times of the year or in some plants or not in others. That ability to distinguish those things in the data I think is going to be very hard because we don't have a nice variable to measure, we don't have something that really varies across plants and across time that's picking up that constraint.

MR. MONCH: I see another question.

MR. MITCHELL: Bob Mitchell. Actually, I'd rather listen to Nalan go on for the rest of the afternoon, but I'm going to follow up on his comment that the plant might not have space for additional machinery.

When we think about volume increases, we often talk about, gee, whiz, what if there's a 10 percent volume increase? I realize this is larger than a first derivative, but at some point we have to talk about a volume increase that's large enough for us to begin to see the actual behavioral characteristics of the plant, so I'll talk about 10 percent.

When we've done transportation analysis over the years, we've often said, well, what's going to

happen to your trucking costs if volume goes up 10 percent and the answer traditionally is, well, on average, our trucks re only 68 percent full, so if we have a 10 percent volume increase, why, most of it we can handle in the existing trucks. But then someone says, you know, some of the trucks are going to get too full and we're going to have to get a whole other truck. So all of a sudden we've got a very low marginal cost in a whole lot of areas, but we've got a very, very big marginal cost in a couple of areas because we had to buy a whole new truck. And so what we wind up doing is kind of averaging this thing out and assuming that we get more trucks and we don't handle this thing as a short-run change in how full the truck is.

Well, if we apply this to mail processing nationwide and I say, well, I've got 500 machines nationwide that I'm using, what's going to happen if I get a 10 percent volume increase?

Well, with a 10 percent volume increase, we might get 50 more machines. Basically, we're going to have to do that to handle our volume.

Where are the 50 machines going to occur? I'm not sure. We build two or three plants a year, we put in some annexes, once we put in a new plant, we

rearrange the volume a little bit, so even though the new plant is over here, we shift volume out of a couple of these plants and we shift volume into the new plant and we essentially accommodate the volume increase and we essentially achieve getting our new equipment.

Well, you know, with a 10 percent volume increase we've got 500 machines and a 10 percent volume increase and we buy 50 more machines somewhere, don't all of the set up costs then become variable? We talk about scheme changes and we talk about set up costs for the machine and so forth. If we start buying new machines, then all the set up costs become variable too and our explanation for why we have variabilities of less than 100 percent tend to disappear.

I guess I'm wondering that if you do anything approaching a longer term or a full adjustment to some of these volume changes, if the fixed scheme changes also become variable.

If this doesn't make any sense, I realize it's a little different from the average question here.

MR. HUME: This is Peter Hume again. This is exactly the question that I felt like asking, too,

but this has come around as far as I know -- Harold Orenstein, who God preserve, I think brought that one up and we used to argue about it and he said everything is 100 percent variable just because you have more of the same.

I think there's a further way of looking at this and that is that the marginal costs enable you to allocate costs among the different classes of mail, which is a different question really, and I don't quite know when you say, well, we've got to add 50 other machines because of the 10 percent volume increase whether that gives you a different insight for allocating the cost differentially to the different operations and, in particular, to the different classes, which is really the problem we're facing. We're not trying to deal with the total cost to the Postal Service, we're trying to do a decent allocation that doesn't annoy people.

MR. BOZZO: This is Tom Bozzo again. In part, my response would be if the existing 500 machines aren't at full capacity, it would seem to be an assumption that may or may not prove true that a 10 percent volume increase would necessarily lead to an increase of 50 machines, on the one hand. I suppose it's also not generally true that all machines -- in

the long run, all machines may be relocatable. In the time between the base year and the test year of a rate case, as a practical matter, most of them are not going to be relocated, so in principle you would expect that volume increases will cause some, probably not many machines, to reach capacity in any particular interval of time. Most will not.

I think as a practical matter when we deal with a postal system with relatively flat overall volume that is generally shifting to more work shared categories, I think contemplating large increase in piece processing volume is perhaps not the most important scenario that we would necessarily want to explore as far as the implications for cost changes.

MR. WALSH: I think I can answer your question more directly as a matter of simple arithmetic. Let's assume all those 500 machines were at full capacity and we have a 10 percent volume increase. Do we have to buy 10 percent more machines? No. And the reason not is simple arithmetic. The reason is because assume that 10 percent of the time on those machines was being used to change the schemes, change the sort plans, relabel the bins and start up a new sort plan.

That 10 percent hasn't changed, it's still

there, we still have all the same destinations, we still have to do all the same sort scheme changes. So you're actually doing that mail with 90 percent of the total time, 10 percent being on scheme changes. That means you have to buy 9 percent more machines, not 10 percent more machines.

MR. THOMAS: Joel Thomas, National Association of Presort of Mailers. When talking about a machine being in capacity, maybe in one shift it is, but there are very few of these machines really running 24/360, are there?

MR. WALSH: That was just a hypothetical. This is Barry Walsh again. That's a valuable point. Because unfortunately the situation we have to deal with nowadays is not so much a 10 percent increase in volume as a prospect of a 10 percent decrease in volume.

Now, nobody has suggested that we scrap a bunch of DBCSs, so we don't really have this problem of buying more machines or less. We do have this fact of life that you have to change the sort schemes every time you get a new set of destinations to sort to and that constant percentage comes out of whatever change you have to make in your total run time.

MR. MITCHELL: Bob Mitchell again. I wonder

whether I can keep this all connected or not. One of the things we have to remember is the volume change and the marginal costs that are relevant to rate making is hypothetical. In other words, it says we are considering one rate instead of another rate in our rate setting process for this particular class of mail and we would like to set that rate in view of the effects of one rate versus another and one of the effects is a cost effect.

So we hypothetically say if the rate was set a little lower instead of a little higher and the volume turned out to be a little higher because of that, given the elasticity, how much would the costs go up?

As economists, we're also stuck with this kind of long-term problem that we've all been trained that the marginal costs for volume increase is exactly the same as marginal costs for volume decrease, that these are continuums and investment agreements a practical matter, if you're in a real organization and the volume goes up next year, you handle it one way and if the volume goes down, you handle it a different way. But we're not here trying to do a roll forward. In other words, a projection of what's going to happen in a test year is really a different question from the

hypothetical question of what if there were a certain kind of volume change. What if the volume change in the test year is a little different form what it was?

So I think we've got to make sure we keep this hypothetical and I think we've got to get away from the question of the roll forward and view that as something completely different.

I would like to be able to constrain ourselves so that the marginal costs of a volume increase is the same as the one for volume decrease. That's a bridge that I've never known quite how to get across, but under the assumption of the tight system, either way, I think it's a little more reasonable to have the same marginal costs for an increase as for a decrease.

MR. HUME: On that comment, I remember one of our brighter guys around, I believe it was Mike Nelson, used what he called the one piece decrement in mail volume delivery in order to estimate a marginal cost and the bright idea was that if you look at the number of single piece stops clearly if you remove one piece then you have removed all those stops, so that the variability of the number of stops depends essentially on the number of one-piece stops and that was the negative removing the volume concepts and I

believe that was quite seriously considered and it was a very good thought and I think you can apply the same thing, what would you remove?

I would like to endorse your point, that we are actually allocating costs among classes of mail and we all seem to agree that marginal costs is a good way to do it and the roll forward is a different question, as you've said. I would agree entirely.

MR. MITCHELL: And I was just thinking while we're all here having fun if we go back to some of the route analyses that Peter was an integral part of in the early 1970s, we always for costing purposes assumed a strict eight-hour day and we assumed that as soon as the volume increased it went over eight hours and we put on more routes and we used to keep the route time the same, the travel time used to go up. If we had a volume decrease in our route analysis, we assumed that the route time went below eight hours and all of a sudden we adjusted all the routes and we did an as-is analysis for volume decrease and we called that a long run adjustment process.

So I think in a trucking analysis and the route analysis, we have this obvious long run adjustment process and variability here. It's looking to me like it's getting shorter and shorter.

MS. DREIFUSS: I see it's 2:00. We've ventured onto a fairly wide ranging discussion which I think it's actually good to come together sometimes and share views like this.

Does anyone want to ask Mark any further questions about his analysis? I see one.

MR. ROBINSON: Dr. Roberts, I'm Charlie Robinson. I'm with the commission. I have a more mundane question to ask and it's about cost allocation.

I'm looking at Table 4 and I look at the four columns there in the first two lines and is authority to myself, okay, suppose I bought off on these output elasticities and I actually wanted to use them for a rate case.

I'm assuming that I would just sum the first two lines in the manual column and apply them to the manual MODS cost pool for letters and I would do the same thing with OCR and the mail processing bar code sorter. And then when I get to DBCS, because I have kind of an accounting mentality, I say to myself, uh-oh, it looks like the number there is 120 percent and I'm going to take 120 percent of the accrued cost for this cost pool and call that an attributable cost.

Well, my brain tilts on that one and it

causes me a problem, so I have two questions: what would your recommendation be to the commission for how to handle the situation where marginal cost exceeds average total cost and, number two, could you intuitively explain to me why you obtained this result? I don't understand why the marginal costs would be so much higher for a delivery bar code system.

Thank you.

MR. ROBERTS: I haven't fully worked this out in my own head, but I think the way I would start is I would not be doing the cost allocation by cost pools like has been done. I would be thinking about shapes of mail and I would be aggregating up to looking at the elasticity by shape of mail.

Let's suppose now we just have one output, that makes it a little bit easier. What I would want to get is the marginal cost, the marginal cost of a letter, and I would work at that level.

I'm not convinced myself and I'm a little bit on thin ice here, so I'm talking to the experts in the room and I've not thought as deeply about this, but I'm uneasy with the need to allocate these costs pool by pool and I think the way I've been thinking about it is allocating them letters, flats, it's

really by shape that that's where the processing costs are coming from.

These are pieces of the marginal cost of letters and we want to take the expenditure on letters and allocate that across classes, so your problem wouldn't come up probably.

MR. ROBINSON: Let me get a clarification. In a simplistic way, are you suggesting that I would just sum up the four cost pools for those letters right there and apply the 100 percent variability factor, do the same for flats and apply a 70 percent?

MR. ROBERTS: Using my estimates, yes. That's kind of the way I'm thinking about it. Yes. And not a separate allocation cost pool by cost pool.

MR. ROBINSON: Does this mean you're not going to attempt to answer why the marginal cost is so high for DBCS?

MR. ROBERTS: Yes. Let me make one point. The allocation by cost pool, that is partly tied up with the separability assumption that's made in the modeling. When you make the separability assumption, then the separate cost pools, process by process by process, make sense. You're dealing with these processes independently. When you give up the separability assumption, all these things become

substitutable inputs in processing letters and so you want to think about the marginal cost and the cost allocation for letters and you don't do it cost pool by cost pool.

Notice what that's also going to imply. I didn't talk at all about the distribution key methodology here, partly because I view that as a complication that I think is making it difficult to understand the real differences in the production model which are different between my approach and the U.S.P.S. approach, so that's why I abstracted from the distribution key to think of just one rate class of mail and then I think it makes the technology assumptions cleaner, that you can see the distinction between it.

But what it also says, you don't need a distribution key by process, you don't need to measure how much of the volume is in each rate class in OCR and in DBCS, you only need to do it for letters and so it simplifies, I think, the whole distribution key approach that you need.

So that view of thinking about separability, separate processes, distribution keys by process, that's all tied up with the assumption of separability in the production model, which I don't think it's as

bad as the proportionality assumption or as harmful as the proportionality assumption, but I still think it's harmful and unnecessary, so I would move away from that.

I realize there are probably a lot of details I'm missing. It's something I want to think about more.

MR. WALSH: Barry Walsh again. If you were to move to your approach, that still will leave the question, though, of how the commission would do its current work with things like the discount models that fuel the drop ship discounts and the like because they are dependent on having these cross pools, the marginal cost reach of the cost pools.

What would you propose as an alternative to that?

MR. ROBERTS: Like I said, I'm only beginning to even think about this problem and so I'm not aware of even what you're talking about, what these things are, so I'm just not qualified to speak on that yet.

MR. WALSH: Then a more general question. You've commented and others have commented that these numbers and these estimates can change rather dramatically depending on the time period that you

select for doing your analysis and in particular in flats or OCRs or something you'll get substantially different results for one set of years compared to another set of years.

Do you think that this is largely a matter of error in the estimate or it's a change in the actual real situation?

MR. ROBERTS: It's probably both. The way I try and look for stability, I can just try and describe what I look at that I find odd and what I look at and I find sensible is when I see estimates that vary a lot when I add one additional quarter or when I subtract ten plants from them, I say something is wrong, this doesn't make sense. Or if I was to estimate the model using 1999 to 2002 and then I add in 2003 and I see things change a lot, that kind of variability bothers me because I think that there's some kind of model misspecification that's probably leading to that.

I do believe this is a production process, it's going to change gradually over time, it should change gradually over time and if I'm not seeing that in the estimates, then I question whether I'm estimating the production technology and that's all I can tell you. Then I go back and I try to think of

things that I left out of the model, what am I missing, what's going on here that I'm missing and sometimes I can find things, sometimes I can't, and the MPBCS is a great example. When you look at those estimates, they're very sensitive to small changes and I don't really believe I'm estimating the technology there very accurately. I just don't think it is.

I also know, look, this operation is being phased in and phased out in these plants and so I know that matters, so I've tried to do some things with the data like if it's the first four quarters in which an operation is used or the last four quarters before it's eliminated, I don't use those operations.

It could be that the MPBCS is something that's not phased out entirely, so I still see it in operation, but the way it's being used in the plant is changing over time as these other technologies are added and that's probably what's responsible for the difference in results and I'm not picking that up in the kind of variables that I have.

What I would like to know is if I had access to people like yourself that I could ask these questions, I'd say what am I missing in the MPBCS? What's going on there in the way that it was used in 1999 versus 2000 versus the way it's used in 2004. I

still see the step there, I realize the hours are less, the TPF is less, but are you using it in a fundamentally different way?

If you came back and said, yes, now we're doing this with it where we used to do that, then I'd think about, okay, is that something I could capture in the model?

The model building here is more of an interactive process, I guess, you know, sort of confronting the estimates with what you know about the technology and then saying does that make sense? And, if not, going back, what am I missing, is it really just bad data?

I think I've been able to make progress over time in identifying things that were missing from the model or ways that I'd improve the model, so I do feel like at least there's progress being made, but if I eliminated everything to the point where there's no specification issues, no, probably not.

MR. WALSH: Just more generally, my understanding is that the ultimate purpose of this is to figure out the elasticity and the marginal costs of the test year, which is few years out, and yet the models that you're working with, of course, are historical in nature and they're looking at what was

the elasticity in some period three or four or five years ago over an average of some years.

MR. ROBERTS: Right.

MR. WALSH: Is there some way to more directly address the question of what it would really be or estimate more directly what it would be in the test year, especially considering the way the estimate seemed to be so heavily dependent on just what period of past years that you chose in the first place?

MR. ROBERTS: One way in which the period matters is that we're estimating elasticities which I'm treating as constants, but that can be generalized but I'm treating them as constants over the historical period that I estimate.

Now, you can still recognize that letter sorting, for example, is a mix of these four operations and the relative importance of these operations changes over time, so even though I have a constant elasticity for each operation, I can recognize that, gee, manual was a lot more important early on and DBCS is much more important now.

So when I'm getting my aggregate, what I call my letter elasticities, those are weighted averages over all the operations. Those weights are labor shares. Those could be specific to any year in

the sample. I've constructed them just at the average over everybody, but I could go into the data and say, no, let's do it using the 2004 shares because they have a much bigger role for DBCS, they have a much bigger role for AFSM, and we know that that's going to be more important two years go. So you could do some adjustments like that. They're not going to make huge differences because these shares don't change drastically from year to year, but they will make some.

So partly it's an estimation question and then partly it's an aggregation question, how do you want to aggregate these estimates up and use them and I think you could think of those a little bit differently.

MR. BOZZO: This is Tom Bozzo and I actually completely agree and disagree with this idea that the roll forward and the base year elasticity estimates are completely separable, that the way the Postal Service has traditionally conceived this is that we estimate costs or elasticities that are somehow representative of the conditions in the base year and then do roll forward type adjustments, including adjusting these cost shares in response to expected changes in equipment deployments over the relevant

time horizon and that's how we wend up getting costs that are a projection of test year conditions.

So the two parts of the analysis are, I think, interrelated in that regard and obviously while there are many things that are different about what we do in this regard, I don't think there's any difference in the underlying principle.

MR. KOETTING: Eric Koetting from the Postal Service. On the cost pools, I'd just like to clarify the question about how would you utilize the cost pools and you said, well, you wouldn't use cost pools, you'd just use the letters but I'm very confused about that.

Is what you're saying that you would use the letter variability and you'd just aggregate all of those cost pools into one letter cost pool? The variability that you have, for example, doesn't apply to allied operations at all, they're not in anything that you've done.

MR. ROBERTS: No, it would just be for letter sorting. It would only be for letter sorting.

MR. KOETTING: So you would still have cost pools, they just wouldn't be the individual operation ones, you would just aggregate letter cost pools and flat cost pools?

MR. ROBERTS: I would aggregate the OCR, DBCS, manual pools into a letter pool. And, in fact, the formula that I present for marginal costs has that underlying them. It's average costs on the right-hand side, it's CL, which is the cost pool for letters and that's the sum over all of these. Yes.

MR. KOETTING: Thank you.

MR. WALSH: Just continuing on that, you know, we will face some of these situations in a few years with the FSS machine which introduces delivery point sequencing for flats and it's a huge investment in new technology and it will surely have its own elasticity and we will face at some point a rate case in which we are projecting a massive buy and installation of these machines over a few years which would change the elasticity and the marginal costs in the test year.

Now you're saying we should do that by assigning different weights.

MR. ROBERTS: If the thing has not been introduced yet, then that's hopeless.

MR. WALSH: No, let's assume it's been introduced so that we have some means of estimating its elasticity.

MR. ROBERTS: Okay. Then you do have some

flexibility to change weights in the way you would aggregate over the sorting steps and if you say, look, we know that we're going to shift our volume toward this particular step, that could be used as an argument for increasing the weight on that step.

MR. WALSH: So in this case, we would have a separate variability for FSS machine as opposed to just a variability for flat sortation in general.

MR. ROBERTS: Okay. So what I'm saying is for the variability on the FSS would be one of the components that would create a variability for flats and the question is how much weight do you give that in creating his aggregate and that's what I was thinking of. You have some flexibility in saying, no, the weight, which is the share of total hours that are in this operation, flat sorting hours in this operation. No, we know that's going to go up over time, so we would argue for having a higher weight than what we see in the current year because we know we're going to increase it.

And it would all just come down to justifying what's the mix of sorting operations that you think are relevant for whatever test year you're supposed to be producing.

MR. WALSH: Doesn't that all presuppose that

you have a separate elasticity measure for each of the cost pools?

MR. ROBERTS: A separate elasticity for each of the sorting operations, yes. Yes. But you don't have to take those cost pools, those sorting operations and allocate costs of sorting operations across rate classes of mail cost pool by cost pool. You do it for flats and you allocate those across rates. You allocate that across rate classes, okay?

MS. DREIFUSS: I want to thank everybody.

I see we have one more and then maybe one more after that, I think.

MR. WALLER: John Waller with the commission. You may have partially addressed this and tried to answer it when you were talking about what happened between flat sorting manual and the AFSM 100. What's the explanation of why when you really are similar type technologies that you have such a difference in variability on letters and flats? What's the basis of that? Are there greater set up times on flats, if that is the cause of it? Is it a mix?

MR. ROBERTS: Yes. I'm going to give you the mechanical answer, John, not the answer you really want, but what it traces back to is -- think of these

operations. Suppose it was just manual and then the newest, fanciest, high tech automated operation, so let's just think of manual and AFSM, manual and DBCS on the letter side.

If we look at the results, we find that the DBCS and the AFSM look sort of similar in their hours response to volume changes. They look roughly the same. If you add the two coefficients up, they come up a little bit more than 1, 1.01 and 1.2 in the other case. The big difference is in the manual and I'm seeing that in letters the manual response is .91 and in flats it's .60 and that's really where all the difference is coming from. You'll notice those shape elasticities that I showed you, so it all traces back to what's going on in the manual operation.

MR. WALLER: Would that warrant some special study, then, of the two manual operations if that's driving it so much?

MR. ROBERTS: Yes, I think that would be very interesting. I've sort of isolated in this data what I can see as the source of the difference. I don't know why that difference is there but I think a case study -- I was hoping someone here maybe would have some ideas on what it was, but, yes, it's got to be matched up with the reality of what's going on in

the plants, not just looking at the data here.

MR. WALLER: They do have different sort rates, no doubt about that.

MR. WALSH: There's some fundamental differences between letters and flats in terms of our service requirements. Letters, you've just got two varieties, basically, your first class and standards, standards deferable. That means for letters all you ever have to sort on the manual cases are things that won't go through the automation, that are rejected or you look at and decide not to put through in the first place.

The flats you have a different situation because you have periodicals and an priority there. They have service requirements. A lot of daily newspapers and the like, they come in way too late to ever go on any machine. They've got to be sorted by hand. A lot of priority, the same thing, you don't have any time to try and put it on a machine. Also, a lot of the sort runs that would be required are much too short. By the time you set the machine up and then sweep it down, you've spent 15, 20 minutes setting up and taking down to run five minutes worth of mail. That doesn't make any sense. This is a fundamental difference between flats and letters and

I suspect that's what behind what you're talking about.

Is that consistent with the nature of the difference you've seen?

MR. ROBERTS: It's even a little more subtle because where I see the change is between this time period and the prior time period that I was working with, the '94 to 2000 data. There, I found the manual variability in flats was .8 something to .9 something -- here it is, .84 to .96. And now I'm getting something like .7. So there's something that's different about the prior time period versus the latter.

Now, of course, the introduction of the AFSM is a big part of it. Thinking the next step, though, why does that change the output variability for manual? It's got to require some kind of change in the way that manual labor is being used in the plant to respond to volume changes. I could imagine there might be some, but I don't know exactly what it might be.

MR. WALSH: The AFSM technology is much more effective than the previously 881 technology and we now have these things fully deployed so we have plenty of capacity.

It used to be that there was a lot more mail that went on manual because the 881s were not anywhere near that much more effective than manual in the first place, so you had a more steady stream of stuff going through there.

Now, all you've got is this high priority stuff, stuff you have to get out. People sit there waiting for it to get there, so you're going to have more waiting time, for example, in the current flats operation than you used.

MR. ROBERTS: There is one fact that is employee in the data that's consistent with that is that this is when you look over time -- in fact, it's Figure 1 in the paper, Figure 1, the bottom part of it, I give you just the total man hours in the four flat sorting operations and you do see that the quarterly variation in hours s quite large in the 881 and the manual in the early part of the time period and as you move into the later part of the time period, the quarterly variation is all picked up by the AFSM and you see the manual has both declines, but it also appears to have less cyclical sensitivity to it, too. I think that's consistent with what you're saying and that's my guess, is that that has some role to play in why these variabilities are less, but

that's a guess.

MS. DREIFUSS: Well, I want to thank everybody today. There was quite a valuable exchange of information and opinion here and very satisfyingly done in a most amicable way possible.

My very special thanks go to Mark Roberts. I think he answered your questions very patiently and I think very thoroughly, certainly to the best of his ability which is considerable.

Thank you all for coming. This is the end of our workshop.

(Whereupon at 2:28 p.m., the workshop in the above-entitled matter was concluded.)

//

//

//

//

//

//

//

//

//

//

REPORTER'S CERTIFICATE

DOCKET NO.: N/A

CASE TITLE: Workshop by Professor Mark Roberts

HEARING DATE: March 14, 2006

LOCATION: Washington, D.C.

I hereby certify that the proceedings and evidence are contained fully and accurately on the tapes and notes reported by me at the hearing in the above case before the Postal Rate Commission.

Date: March 14, 2006

Christina Chesley
Official Reporter
Heritage Reporting Corporation
Suite 600
1220 L Street, N.W.
Washington, D.C. 20005-4018