Postal Regulatory Commission
Submitted 9/23/2010 2:55:25 PM
Filing ID: 70271
Accepted 9/23/2010

USPS-RT-1

BEFORE THE
POSTAL REGULATORY COMMISSION
WASHINGTON, DC 20260-0001

SIX-DAY TO FIVE-DAY STREET DELIVERY

AND RELATED SERVICE CHANGES, 2010

Docket No. N2010-1

REBUTTAL TESTIMONY

OF

DR. PETER BOATWRIGHT

ON BEHALF OF THE

UNITED STATES POSTAL SERVICE

TABLE OF CONTENTS

Biographical Sketch iii
Purpose of Testimony 1
Executive Summary 3
Context: Product Decisions and Uncertainty 4
Specific Issues Raised by NALC Witness Crew 7
Use of likelihood scale 8
Potential bias 13
Alternative or Additional Analyses 20
ORC Analysis Is Not Dominated by Alternatives 20
Potential Benefit of Additional Research and Analysis is Dwarfed by Cost of Delay 25
Concluding Remarks 27

List of Tables:

1. Gross and Net Savings for Five-Day Delivery 2
2. Examples of Residual Bias, with Minimal Impact on Financial Benefit of Five-Day Delivery Proposal 20
3. Example of a Discrete Choice Question 24

BIOGRAPHICAL SKETCH

My name is Peter Boatwright. I am an Associate Professor of Marketing in the Tepper School of Business at Carnegie Mellon University, where I have been on the faculty since 1997. I have an M.S. in Statistics from the University of Wisconsin, Madison. My M.B.A. and Ph.D. are both from the University of Chicago's Booth School of Business.

My research focuses on product development methods, consumer response models, and Bayesian econometric modeling. In the area of product development, for example, I have formulated methods to assess customer needs and opportunities, and to translate those into product specifications early in the design process. I have also developed statistical models of consumer demand for products. My publications are found in the *Journal of the American Statistical Association, Marketing Science, Management Science, Journal of Marketing Research, Decision Analysis, Journal of Mechanical Design, Research in Engineering Design, Journal of Marketing,* and *Harvard Business Review*. Following is a list of my articles published in refereed journals:

Boatwright, Peter, Sharad Borle, and Joseph B. Kadane (2010) "Common Value/Private Value Categories in Online Auctions: A Distinction without a Difference," *Decision Analysis*, Vol. 7, No. 1, March, pp. 86-98.

Orsborn, Seth, Jonathan Cagan, and Peter Boatwright (2009) "Quantifying Aesthetic Form Preference in a Utility Function," *Journal of Mechanical Design*, Vol. 131, 061001-1-10.

Boatwright, Peter, Jonathan Cagan, Dee Kapur, and Al Saltiel (2009) "A Step by Step Process to Build Valued Brands", *Journal of Product and Brand Management*, 18, 1, 38-49.

Orsborn, Seth, Peter Boatwright and Jonathan Cagan (2008) "Identifying Product Shape Relationships Using Principal Component Analysis" *Research in Engineering Design*, 18, 4 (Jan), 163-180.

Orsborn, Seth, Jonathan Cagan, and Peter Boatwright (2008) "A Methodology For Creating A Statistically Derived Shape Grammar Composed Of Non-Obvious Shape Chunk", *Research in Engineering Design*, 18, 4 (Jan), 181-196.

Lu, Jiang, Joseph B. Kadane, and Peter Boatwright (2008) "The Dirt on Bikes: An Illustration of CART Models for Brand Differentiation", *Journal of Product and Brand Management*, 17, 5, 317-326.

Boatwright, Peter, Ajay Kalra, and Wei Zhang (2008) "Research Note: Should Consumers Use The Halo To Form Product Evaluations?" *Management Science*, 54, 1 (Jan), 217-223.

Boatwright, Peter, Suman Basuroy, and Wagner Kamakura (2007) "Reviewing the Reviewers: The Impact of Individual Film Critics on Box Office Performance", *Quantitative Marketing and Economics,* 5, 4 (Dec.), 401-425.

Boatwright, Peter, Sharad Borle, and Kirthi Kalyanam (2007), "Deconstructing Each Item's Category Contribution," *Marketing Science*, 26, 3 (May-June), 1-15.

Borle, Sharad, Peter Boatwright, and Joseph B. Kadane (2006) "The Private/Common-Value Continuum: An Empirical Investigation Using eBay Online Auctions", *Statistical Science*, 21, 2 (May), 194-205.

Kamakura, Wagner, Suman Basuroy, and Peter Boatwright (2006) "Is Silence Golden? An Inquiry into the Meaning of Silence in Professional Product Evaluations," *Quantitative Marketing and Economics*, 4, 2 (June), 119-141.

Kadane, Joseph B., Galit Shmueli, Tom Minka, Sharad Borle, and Peter Boatwright (2006) "Conjugate Analysis of the Conway-Maxwell-Poisson Distribution," *Bayesian Analysis*, 1, 2, 363-374.

P.B. Seetharaman, S. Chib, A. Ainslie, P. Boatwright, T. Chan, S. Gupta, N. Mehta, V. Rao, and A. Strijnev (2005) "Models of Multi-Category Choice Behavior." *Marketing Letters*, Volume 16, #3.

Borle, Sharad, Peter Boatwright, Joseph Nunes, Joseph Kadane, and Galit Shmueli (2005) "Effect of Product Assortment Changes on Consumer Retention," *Marketing Science*, Vol 24, No. 4 (Fall), pp. 616-622.

Shmueli, Galit, Minka, Thomas P., Kadane, Joseph B., Borle, Sharad and Boatwright, Peter (2005), "A Useful Distribution for Fitting Discrete Data:

Revival of the COM-Poisson", *Journal of the Royal Statistical Society, Series C*, vol. 54 (1), pp. 127-142.

Boatwright, Peter, Sanjay Dhar, and Peter Rossi (2004), "The Role of Retail Competition, Demographics, and Account Retail Strategy as Drivers of Promotions Sensitivity," *Quantitative Marketing and Economics*, 2, 169-190.

Telang, Rahul, Peter Boatwright, and Tridas Mukhopadhyay (2004) "A Mixture Model for Internet Search Engine Visits," *Journal of Marketing Research*, 41, 2 (May), 206-214.

Boatwright, Peter, Sharad Borle, and Joseph Kadane (2003) "A Model of the Joint Distribution of Purchase Quantity and Timing," *Journal of American Statistical Association*, 98, 463, 564-572.

Lee, Jonathan, Peter Boatwright, and Wagner Kamakura (2003) "A Bayesian Model for Prelaunch Sales Forecasting of Recorded Music," *Management Science*, 49, 2 (February), 179-196.

Nunes, Joseph and Peter Boatwright (2001) "Incidental Prices and Their Effect on Consumer Willingness to Pay", *Journal of Marketing Research*, 41, 4, 457-466.

Boatwright, Peter, and Joseph Nunes (2001), "Reducing Assortment: An Attribute-Based Approach," *Journal of Marketing*, 65 (July), 50-63.

Boatwright, Peter, Robert McCulloch, and Peter Rossi, (1999) "Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model," *Journal of the American Statistical Association*, 94, 448, 1063-1073.

I have also published two business books on the topic of innovation: *Design of Things to Come*, and *Built to Love*.

I have worked with many companies pursuing innovation in their new product development, including Apple, Bayer Materials, Nissan, MSA, GlaxoSmithKline, among others.

PURPOSE OF TESTIMONY

The Postal Service has proposed reducing the delivery of mail from the current six days per week to five days per week to street addresses (as is explained in the testimony of witness Samuel Pulcrano, USPS-T-1), referred to in this testimony as the "Five-Day Delivery" proposal. The positive financial impact of this proposal has been estimated to be about $3.1 billion (annually USPS-T-7, attachment 3, page 2), or 4.5 percent of Postal Service operating revenues.[1] The Postal Service's loss from operations in 2009 was $3.7 billion[2], meaning that the Five-Day Delivery proposal would have significantly decreased that year's loss if it had been implemented.

The purpose of my testimony is to evaluate the estimates of the volume, revenue and contribution portion of the calculated $3.1 billion annual financial benefit from moving to Five-Day Delivery. Although the estimation of the annual benefit entails numerous calculations (see for example USPS-T-6), the estimation process can be summarized as cost savings less contribution loss.

In Table 1, the cost savings is estimated to be $3.3 billion (Line A). For the purpose of my testimony, I assume that this estimate of the gross annual (cost) savings, $3.3 billion (Line A), is correct. As reflected in the oral cross-examination of

[1] Operating revenues used are reported in the Postal Service annual report for 2009, http://www.usps.com/financials/_pdf/annual_report_2009.pdf. The net financial savings in USPS-T-7 also uses FY 2009 figures.

[2] Postal Service annual report for 2009, http://www.usps.com/financials/_pdf/annual_report_2009.pdf

witness Crew (Tr. VIII/2523, lines 6-11), witness Crew accepts the $3.3 billion figure as well.

Table 1--Gross and Net Savings for Five-Day Delivery

In FY 2009 Terms

Replicated from USPS-T-7, attachment 3, page 2 and Revised USPS-T-9, pp. 14-15

(in millions of dollars)

Line A	Gross Savings	$	3,300
Line B	Revenue Loss	$	(466)
Line C	Cost Reduction	$	260
Line D	Contribution Loss	$	(206)
Line E	Net Savings	$	3,094

Continuing with Table 1, the contribution loss due to reduced mail volumes has been estimated to be $206 million (Line D, from revised testimony of United States Postal Service witness Gregory M. Whiteman, USPS-T-9, pp. 14-15). The contribution loss estimate reflects a volume loss estimate of 0.71 percent based on the Opinion Research Corporation (ORC) quantitative market research. The contribution loss is a sum of revenue loss of $466 million (Line B) and a cost reduction of $260 million (Line C), where the cost reduction occurs as a result of the lower volume of mail to process and deliver.

My testimony focuses on the reliability of the estimates of the mail volume reduction that would be triggered by Five-Day Delivery, the portion of the calculation that leads to the contribution loss of $206 million (Line D).[3]

To the extent that the gross annual cost savings estimate of $3.3 billion is not in question, the decision centers on the estimate of the contribution loss, $206 million. If the contribution loss is truly near $200 million, then the financial benefit (*i.e.*, the net savings) of the Five-Day Delivery proposal will be $3.1 billion. The word "near" in the prior sentence is the crucial word. How near is the true value to the estimate of $206 million? For example, if "near" means somewhere in the range of $100 million to $300 million, financial benefit of the Five-Day Deliver proposal will exceed $3 billion. This word "near" therefore can be thought of as summarizing the discussion surrounding the reliability and validity of estimates developed by the Postal Service based largely upon the market research conducted by ORC. To address this question, we need to assess the reliability of the volume loss estimate of 0.71 percent.

EXECUTIVE SUMMARY

Based on an evaluation of the ORC research methodology and other witness testimony, I view the estimate of a contribution loss of approximately $200M to be a reasonable estimate (Line D in Table 1). Added to the $3.3B gross cost savings estimate (Line A in Table 1), I conclude that the Five-Day Delivery proposal will generate an estimated annual savings of $3.1B for the Postal Service.

[3] Numerical figures in this paragraph were taken from USPS-T-7, attachment 3, page 2.

I see no reason why additional research and analyses would meaningfully change the estimates already in hand. Also, additional research and analyses would be expensive in itself especially due to a delay of implementation of the Five-Day Delivery proposal, for a delay could easily cost over $1.5B.

After considering the market research, analysis, and the financial estimates Postal Service experts developed based on the ORC research and analysis, it is my conclusion that the Five-Day Delivery proposal offers compelling financial value.

CONTEXT: PRODUCT DECISIONS AND UNCERTAINTY

An evaluation of the financial consequences and whether to implement the Five-Day Delivery proposal is analogous to that of a firm considering whether to change one of its products. In the field of marketing, changed products are considered "new" products, where "new" denotes a potentially wide range of changes, from marginal changes in existing products to launches of products that are significantly different from others in a firm's current portfolio. For the firm, a new product has financial implications for the future of the organization that are critical to what I will call "the new product decision," whether to continue the firm's plans to develop, design and launch the new product.

Predicting the future involves some degree of uncertainty, including prediction of future costs, volumes, and revenues. To reduce the uncertainty surrounding a new product decision, firms conduct market research, forecast sales volumes, and analyze projected costs. Critically important, the purpose of gathering

and analyzing data is not to eliminate uncertainty but rather to reduce uncertainty, because in the end the future always retains some uncertainty.

When making decisions about changes to existing product or launching new products, market research is often conducted to support the basic "go" / "no-go" decision. I use the word "product" in its general sense, including services, software, systems, and technologies. Customer views constitute critical input to product decisions, so both qualitative and quantitative marketing research methods are typically employed to obtain customers' perspectives.[4]

Goals of qualitative research include the identification and in-depth understanding of specific customer needs and desires. Popular methods of qualitative research are focus groups and individual in-depth interviews, methods that may result in a breadth of insights about a product change that can later be explored more precisely with quantitative research. Qualitative research is also often used to assist in the design of the quantitative research and to test statements describing the new product.

Although many different types of quantitative analyses, corresponding to various purposes, can be undertaken, the general goal of quantitative work is to use results from a sample of respondents to make inferences about the population (consumers and/or businesses).

Many different types of information are provided to potential respondents in qualitative and quantitative market research that help them make more precise and

[4] Urban, Glen L. and John R. Hauser (1993) *Design and Marketing of New Products*, Upper Saddle River: Prentice-Hall, pp. 182-85.

realistic assessments of likely future product usage. Once a product is launched into the marketplace, customers will have much information available and accessible, any of which may affect purchase decisions. Common market research techniques used to provide the information consumers would have available include:

1) usage scenarios, brief stories that illustrate how a product is intended to benefit the customer thereby providing a context for a buying decision and possible reasons for purchase;
2) product descriptions that provide details which allow customers to evaluate the product more carefully and respond to specific attributes; and
3) where relevant, customers may be given potential marketing details such as price, advertising copy, and promotions.

These techniques help respondents more realistically assess their own purchase likelihood in light of the information that would be available in the marketplace.

The information provided in the ORC customer research is typical of studies conducted by businesses worldwide for product decisions. Given that the USPS products (*i.e.,* First-Class Mail, Express Mail, Priority Mail, Periodicals Mail, Standard Mail, Parcel Select and Parcel Post) are already known to customers, study participants have less need for full product descriptions. Rather, study participants could be given information about the proposed product changes (*e.g.* changes to delivery days) as well as information pertinent to those changes (*e.g.,* Post Offices will remain open) or other contextual reasons for the change (*e.g.,* mail volumes have decreased). Respondents in the ORC research were also given a context for

product changes. The text of this context was extensively tested and refined during the qualitative research phase. This context provided them a scenario in which to consider their expected product usage. As such, the information provided to study participants by ORC follows marketing research standards designed to provide information needed by study participants to make informed decisions.

SPECIFIC ISSUES RAISED BY NALC WITNESS CREW

Below I analyze and explain the weakness or inapplicability of three criticisms lodged by National Association of Letter Carriers (NALC) witness Crew. The first criticism, ORC's use of a "likelihood scale," (NALC-T-4, pp. 5-7) fails to acknowledge or recognize that a "likelihood scale" is a market research industry standard tool supported by an extensive body of peer-reviewed literature. ORC's use of a likelihood scale is fully consistent with that body of knowledge and standard statistical approaches. Not using a likelihood or similar scale would result in biased results.

Next, I consider witness Crew's assertion that ORC's research approach introduces bias (NALC-T-4, pp. 4-5). ORC's research design overtly recognizes specific potential sources of bias common to customer research and uses industry standard approaches to constrain or eliminate bias.

Last, I examine the potential utility of other research or analytical methods that witness Crew suggests would have been better approaches, econometric analysis and discrete choice modeling. I conclude that while such methods may be useful for further reducing the uncertainty inherent in the decision, the cost of the

delay from using such methods now would lead to an expense that is an order of magnitude greater (in billions of dollars) than any financial benefit of the new information (in millions of dollars).

USE OF LIKELIHOOD SCALE

ORC uses a "likelihood scale" in its analysis of Five-Day Delivery's impact on mail volume. The ORC likelihood scale measured purchase intent. Purchase intent measures (measures of the likelihood that customers will purchase a given product or service or, in this case, likelihood that customers would change their behavior if Saturday street delivery is eliminated) are routinely used in market research,[5] for forecasting in general[6] and for new product testing.[7]

Purchase intentions are commonly collected using verbal scales or with combinations of verbal and probability scales. In verbal scales, answers are grouped into categories such as "yes, definitely will buy," "yes, probably will buy", and so forth.[8] A combination of verbal descriptions and numerical probabilities is also used in market research. The Juster scale, for example, is an 11 point scale with categories such as "Certain, practically certain (99 in 100)" "Almost sure (9 in 10)" towards one end of the scale and "Very slight possibility (1 in 10)" "No chance,

---

[5] Morwitz, Vicki G., and Schmittlein, David. (1992). "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which "Intenders" Actually Buy?" *Journal of Marketing Research*, 29(4), 391-405.

[6] Infosino, Willliam (1986). "Forecasting New Product Sales From Likelihood of Purchase Ratings," *Marketing Science*, 5 (Fall), 372-84.

[7] Silk, Alvin J. and Glen L. Urban (1978), "Pre-Test-Market Evaluation of New Product Goods: A Model and Measurement Methodology," *Journal of Marketing Research*, 15 (May), 171-91.

[8] Clawson, C. Joseph (1971), "How Useful Are 90-Day Purchase Probabilities?" *Journal of Marketing,* Vol. 35, No. 4 (Oct., 1971), pp. 43-47.

almost no chance (1 in 100)" towards the other end of the scale.[9,10] ORC used an 11-point version of the verbal purchase intention scale to assess the likelihood that users would change their mailing volume with the USPS (see USPS-T-8, p. 12 and Appendices F and G). The Juster scale is very commonly used in market research and its use has been well studied and documented in the academic literature.[11]

In economics, use of the Juster scale is not common (see question from Chairman Goldway (Tr.VIII/2586, line 19)). However, some leading econometricians are arguing for greater use of approaches like it. Northwestern University Professor Manski has written a review article on the use of subjective probabilities in the field of economics. After showing problems with the typical approach, he writes

> If choice data alone do not suffice to infer how persons make decisions with partial information, one might anticipate that economists would ask persons about their preferences and expectations. However, economists have been deeply skeptical of subjective statements; they often assert that one should believe only what people do, not what they say. As a result, the profession for many years enforced something of a prohibition on the collection of subjective data.[12]

After giving credit to Juster (1966) as the originator (for the field of economics) of a scale measuring people's expectations in order to predict behavior, Manski reviews past work and highlights recent research that has begun to

---

[9] Urban, Glen L. and John R. Hauser (1993) *Design and Marketing of New Products*, Upper Saddle River: Prentice-Hall, p. 307.

[10] Juster, F. Thomas (1966) "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design," *Journal of the American Statistical Association*, Vol. 61, No. 315 (Sep.), pp. 658-696.

[11] Brennan, Mike (2004) "The Juster Purchase Probability Scale: A Bibliography," *Marketing Bulletin*, http://marketing-bulletin.massey.ac.nz

[12] Manski, Charles F. (2004) "Measuring Expectations," *Econometrica*, Vol. 72, No. 5 (Sep.), p. 1337.

1 incorporate purchase intentions into econometric modeling. For marketers, the use
2 of any of the scales above–including the one used by ORC here–is common.
3 It has been stated that "the single best predictor of an individual's behavior
4 will be a measure of his intention to perform that behavior."[13] On both theoretical
5 and empirical grounds, intention measures are believed to improve estimates of
6 future actions.[14] Some benefits of purchase intentions are that they allow
7 respondents to assess their own behavior after independently considering possible
8 factors that may contribute to a purchase decision, including factors that are
9 dynamic. Use of purchase intentions is prevalent for new product research to the
10 extent that "the buying intention question appears in almost every concept test."[15]
11 Witness Crew notes (NALC-T-4, pp. 5-6) that ORC used the likelihood factor
12 to adjust the volume estimates given by customers, resulting in a question
13 concerning how to interpret ORC's "adjusted volume." (*See* USPS-T-8, p. 36)
14 ORC's use of the likelihood factor can be explained definitively using a
15 common statistics equation. The "expected value," also called the "arithmetic
16 mean," "mean," or "average," is a fundamental concept taught in statistics classes
17 and is almost certainly the most prevalent statistic used for any estimation,
18 including forecasts. In many contexts, it may be calculated as simply the sum of the
19 outcomes divided by the total number. For example, to calculate the expected value

---

[13] Fishbein, M. and I. Ajzen (1975), *Belief, Attitude, Intention, and Behavior.* Reading, MA: Addison-Wesley Publishing Company, quoted on page 391 of Morwitz and Schmittlein (1992).

[14] Manski, Charles F. (2004) "Measuring Expectations," *Econometrica*, Vol. 72, No. 5 (Sep.), pp. 1329-1376.

[15] Crawford, Merle and Anthony Di Benedetto (2008) *New Products Management*. Singapore: McGraw-Hill, p. 197.

of a roll of a standard die, one would sum the numbers 1 through 6 (because those are the numbers on the sides of the die), then divide by 6 (because the die has 6 sides).

Notice that in the above example, each outcome (*i.e.*, number on the die) is equally likely to appear. When outcomes are not equally likely, the expected value formula explicitly accounts for their respective probabilities. The expected value is often described as the average value of a random experiment that is conducted a large number of times. For many, it might be most easily understood through use of an example. Consider a situation in which a person rolls a standard die. If a "3" is rolled on the die, the person receives $100. If any other number is rolled, the person receives $1. In this example, the person has 1/6 probability of receiving $100 and 5/6 probability of receiving $1. If a large number of individuals play this game, say 1200 people, the expected value is the average amount each receives. If 1/6 of the people (which is 200 people) received $100, and 5/6 of the people (1000 people) receive $1, then the average payout is ($100*200 people+ $1*1000 people)/1200 people = $16.83. Or, the same formula rewritten is (1/6)*$100 + (5/6)*$1 = $16.83.

This formula, generalized, is the formula for expected value. If there are two outcomes A and B, and the probability of outcome A is p_A, and the probability of B is p_B, then the expected value of these outcomes is A*p_A + B*p_B. Whether recognized or not, it is the general version of the formula used every time someone calculates an average. This formula is taught in statistics classes throughout the world, and it is, without any doubt in my mind, the most prevalent statistic used in forecasting.

ORC asked respondents the following question:

> "If the service change described does happen, what is the likelihood that this change would cause your organization to modify the number of individual pieces of mail and/or packages your organization sends or the way you send it?" (USPS-T-8, Appendix F p. 104.)

This question concerns a probability, the probability that a respondent's mail volume will be affected by implementation of the Five-Day Delivery proposal. If this probability is p, then 1-p is the probability that a respondent's usage of USPS will be unaffected by the proposal's implementation. These probabilities, along with respondents' volume usage estimates, allow us to calculate the expected value of USPS volume by respondent:

$$\text{Expected value of future USPS volume} = p * V_1 + (1-p) * V_2, \qquad (1)$$

where

V_1 is the volume of mail given that the Five-Day proposal implementation **will affect** mail volumes, and

V_2 is the volume of mail given that the implementation of the Five-Day proposal **will not affect** mail volumes.

Now that we have established the expected value, we can compare this calculation to that of ORC in order to assess whether they overestimated or underestimated future volume. On page 31 of USPS-T-8, witness Elmore-Yalch stated ORC's equation as "The difference between the next 12-month volume and the volume in the first 12 months after institution of five-day delivery was multiplied by this likelihood of change percentage, and the result was added back to next 12-month volume." Note that "the next 12-month volume" in USPS-T-8 is the ORC estimate of V_2 above, and that "the volume in the first 12 months after institution of five-day delivery" is the ORC estimate of V_1 above. Using the same notation as above, the ORC formula can be written as

$$p*(V_1 - V_2) + V_2 \qquad (2)$$

where p is the likelihood of change probability. This formula can be re-written as

$$p*V_1 - p*V_2 + V_2 \quad (3)$$

or

$$p*V_1 + (1-p)* V_2 \quad (4)$$

Note this equation (4) is identical to the expected value equation (1). As stated before, the expected value is the statistic most commonly used for forecasts. Apparently witness Crew did not recognize that ORC used the expected value when stating that "As far as I can recall, I have never seen anything like this, and I believe it is a serious flaw." This may be because it was not stated as the formula I use above. I note that witness Crew is very familiar with the expected value formula, as shown in his own work (*e.g.*, Crew and Kleindorfer, 1976).[16]

The calculations used by ORC are identical to the equation for expected value, and thus ORC has used the most well accepted approach for estimation in the field of statistics.

POTENTIAL BIAS

The USPS has invested time and money to assess both the benefits and the costs of Five-Day Delivery. Every new product proposal entails some degree of uncertainty; this includes the USPS Five-Day Delivery proposal. Witness Crew expressed concern about this uncertainty in stating that the surveys such as that performed by ORC may be subject to significant bias (NALC-T-4, pp. 4-5). Here I

[16] Crew, Michael A. and Paul R. Kleindorfer (1976) "Peak Load Pricing with a Diverse Technology," *The Bell Journal of Economics*, Vol. 7, No. 1 (Spring, 1976), pp. 207-231.

consider how potential bias might influence the decision regarding Five-Day Delivery and how ORC considered these potential biases.

It is tautological that survey research entails many potential sources of bias, often referred to in the market research literature as measurement error. The literature on survey bias is not restricted to the academic realm but is well known to practitioners as well. Lists of potential sources of survey bias as well as approaches to address such bias may be found, for example, in market research textbooks as well as in peer reviewed literature.

For this and other reasons, companies hire professional market research firms that are well-versed about measurement error (bias) and accustomed to crafting surveys to avoid it.

A general procedure in such research is to identify sources of uncertainty, identify appropriate strategies that may help reduce that uncertainty, and analyze gathered information with the intent of reducing the uncertainty inherent in all business decisions—thereby improving the value of all information collected to inform the decision being considered. ORC took this exact approach to address potential bias; I give several examples taken from the testimony of witnesses Whiteman and Elmore-Yalch.

One potential source of bias, aggregation bias, would occur when heterogeneous markets are treated as homogeneous. As noted in USPS-T-8, page 30, ORC and the Postal Service identified five market segments, thus identifying this as a potential source of bias. ORC gathered information on each individual segment,

and analyzed the data by segment and product to obtain separate estimates for each product and segment.

Another source of bias stems from the availability of information to the respondent when asked to estimate future purchase behavior. For example, if the product concept described in a survey significantly differs from the actual future product, respondents' estimates of their purchase behavior may be biased. As noted in USPS-T-9, pp. 15-16, the information made available to respondents about Five-Day Delivery differs from the current proposal, in that it was more restrictive than the operational concept currently proposed by the USPS. As such, ORC and USPS identified this potential source of bias and properly concluded that the research volume loss estimates would likely exceed volume loss estimates based on the current less restrictive proposal.

Similarly, if certain information respondents deem relevant to their future purchase context is not provided, they will make their own assumptions that can lead to "attribution bias." The degree to which assumptions are inaccurate can affect forecasts and bias results. For example, the reasoning behind Five-Day Delivery can be relevant to purchase behavior. Research has shown that when incomplete information is provided about a change, people are likely to supply their own inferences about why a change is occurring, thereby affecting their decisions and judgments.[17] An example of ORC's attempt to constrain potential attribution bias is found in the survey where ORC informed respondents that the Postal Service faces major financial challenges.

[17] Gilbert, D. T., and Malone, P. S. (1995). The Correspondence Bias. *Psychological Bulletin*, 117, 21-38.

1 Witness Crew specifically addressed two additional areas of bias. The first is
2 that respondents may bias their responses in an effort to please the interviewer,
3 called social desirability bias. Tr. VIII/2514-15. Dr. Crew is correct in that this does
4 occur. This bias is less when surveys involve greater degrees of anonymity (*i.e.,* use
5 of telephone surveys instead of face to face interviews).[18] A strategy commonly
6 employed to address this bias is to decompose the questions – that is, asking
7 respondents to provide behavioral frequency responses to a series of questions
8 rather than a single overarching question. In a sociology study examining the
9 frequency and quantity of alcohol consumption, asking respondents to provide data
10 on consumption across 12 different drinking establishments indicated dramatic
11 increases in the frequency and amount of consumption than when asked about
12 consumption across a single month.[19]
13 Market researchers also recognize this potential bias when asking
14 respondents to provide behavioral frequency responses or volume estimates to
15 questions. Research has shown, for example, that employing *decomposition strategy*
16 (unbundling a total into subcategories) when asking customers to estimate past and
17 future credit card purchases helped customers recall their individual expenses,
18 thereby reducing sampling-based bias.[20] ORC did exactly this by asking respondents
19 to provide estimates of volume by application (i.e., payments, correspondence, etc.)

---

[18] Richman, Wendy L., Suzanne Weisband, Sara Kiesler, and Fritz Drasgow (1999) "A Meta-Analytic Study of Social Desirability Distortion in Computer-Aided Questionnaires, Traditional Questionnaires, and Interviews," *Jounal of Applied Psychology,* 84, 5, 754-775.

[19] Mooney, Linda A. and Gramling, Robert, "Asking Threatening Questions and Situational Framing: The Effects of Decomposing Survey Items," *The Sociological Quarterly*, Vol. 32, No. 2 (Summer, 1991), pp. 289-300

[20] Srivastava, Joydeep and Raghubir, Raghubir, "Debiasing Using Decomposition The Case of Memory-Based Credit Card Expense Estimates," *Journal of Consumer Psychology*, *12*(3), 253–264.

and then allocating those changes across products (*i.e.*, First-Class Mail, Priority Mail, Express Mail, *etc.*) rather than asking respondents to provide an overall estimate of impact for their business or household.

The final possible source of bias suggested by witness Crew concerns the time interval between when the research was conducted and implementation of the change because of the possibility that the environment may be very different in the future. This is always the case in forecasting and was addressed earlier in my testimony when I noted that there is uncertainty in most decisions. Clearly, the estimates are based on respondents' understanding of the proposed changes to their service in the current environment.

The objectives of the quantitative research were to estimate the change in volume resulting from a change to service in the current environment and to present to the PRC testimony within this environment to make a "go" / "no-go" decision, that is discontinue or continue six-day delivery of mail. This situation is analogous to product introduction decisions. By way of example, consider the pharmaceutical industry, where the time line between the initial decision to "go" / "no-go" and the actual introduction of a new drug involves years. To determine potential future demand, the pharmaceutical industry analyzes data that are concurrent with the "go" / "no-go" decision. The pharmaceutical industry also conducts extensive research with experts and consumers to estimate the need for a pharmaceutical product that offers a new treatment. Based on this information, a specific company within the industry may make the decision to invest the millions of dollars necessary to develop and test the drug. As the time for product introduction

1 approaches, these companies often conduct additional secondary and market
2 research to determine if there are any changes in disease prevalence or physician
3 and consumer behaviors that could affect their original estimates.
4 In the present case, the Postal Service has assembled data to assist in making
5 an informed decision as to whether eliminating Saturday street delivery is an
6 appropriate strategy to address its financial challenges. This is a classic business
7 decision – should the Postal Service eliminate / not eliminate six-day delivery
8 service. Depending on the length of time between this go / no-go decision and its
9 implementation, a meaningful change in the marketing and operating environment
10 may require further analysis and even new research to determine whether the
11 projected impacts have varied enough such that one should consider modifying the
12 operational strategies to mitigate any impacts. It is important to note that based on
13 the analysis done here, any such change would have to be very substantial to
14 warrant new research.
15 The potential for bias in survey research is well recognized in practice and
16 provides a sound argument for using professional research firms such as ORC to
17 design and implement such surveys. The methods used by ORC are consistent with
18 those used in practice by other professional market research firms; the research
19 protocol, procedures and materials are designed to address potential biases that
20 exist in such studies. Properly designed research makes the results reliable, all of
21 which explains why market research firms' results routinely inform the critical
22 business decisions of those who hire them.

Because the Postal Service's ultimate goals in studying Five-Day Delivery included both refining its proposal and deciding whether to move forward with that proposal—a major financial decision—the implications of any residual bias can also be considered usefully in the context of that decision. We can suppose bias does exist—in both directions, quantify it and then inquire whether its magnitude should impact the ultimate financial picture. In other words, what if, even after ORC has addressed well known biases, there remains some residual unrecognized bias that leads to errors of a magnitude of +/- 1 percent, or +/- 2 percent, or even +/- 5 percent? Table 2 illustrates how potential bias of various magnitudes impacts the overall financial benefit of the Five-Day Delivery proposal.

Table 2 – Examples of Residual Bias, with Minimal Impact on Financial Benefit of Five-Day Delivery Proposal

Example Bias Levels	Contribution Loss	Financial Benefit
-5%	$196M	$3.1B
-2%	$202M	$3.1B
-1%	$204M	$3.1B
USPS estimate	$206M	$3.1B
+1%	$208M	$3.1B
+2%	$210M	$3.1B
+5%	$216M	$3.1B

Table 2 shows that even if bias approaches five percent, positive or negative, it has essentially no impact upon the Postal Service's calculation of the financial benefits it calculates Five-Day Delivery would bring. The financial benefit remains at about $3.1 billion. Although amounts in millions of dollars are certainly substantial, any possible impact of unconstrained bias would be dwarfed by the

overall cost savings. So, if the Five-Day Delivery decision were to be based on net cost savings alone, potential biases of these levels would seem unlikely to alter the decision.

ALTERNATIVE OR ADDITIONAL ANALYSES

ORC Analysis Is Not Dominated by Alternatives

Witness Crew notes that other research approaches could have been used to estimate mail volumes for Five-Day Delivery. Two specific approaches were brought up as alternatives, econometric and time series analysis (NALC-T-4, p. 7) and discrete choice models (Response to USPS/NALC-T4-17). I discuss each of these below, after clarifying that the ORC analytical approach embodies the industry standard, not an inferior alternative; ORC's general approach has even been found superior to model-based approaches in certain estimation problems similar to the one at hand here.

ORC Used a Standard and Reasonable Estimation Method

Forecasts of product sales volumes are important for firms' new product decisions, as such estimates have direct implications on profitability for a new venture or product change. Because predicting the future necessarily involves some degree of uncertainty, the purpose of gathering and analyzing data is not to eliminate uncertainty but simply to reduce it.

One task is to inform respondents sufficiently so that they consider a purchase context that is close to what the actual purchase context entails. As a guiding principle, customers should be informed of factors that influence their decisions. Such information will naturally vary depending on context such as what

respondents already understand about the new product. For example, respondents may be given usage or purchase scenarios (*i.e.,* descriptions of a purchase context), product features, and availability of competitive alternatives, as appropriate to the context.

The ORC research is typical of study conducted by industry for new product decisions. Given that the USPS products are already known to customers (*i.e.,* First-Class Mail, Express Mail, Priority Mail, Periodicals Mail, Standard Mail, Parcel Select and Parcel Post), respondents were reminded of those products and told of a usage scenario that provides context for their usage intent. After being given relevant information, respondents were requested to forecast their mail volume, a standard approach for obtaining volume estimates.

Why Econometric and Time Series Models Are Not Best

For some forecasting problems (vehicle sales in France, vehicle sales in the U.S., and U.S. wireless services), researchers documented that sales forecasts based on purchase intentions were more accurate than forecasts based on extrapolations of past sales.[21] So while it may seem at first glance that econometric models of empirically observed data may improve purchase intention data such as that gathered by ORC, the elegance of econometric models does not always translate into

[21] Armstrong, J Scott, Vicki G Morwitz, & V Kumar (2000) "Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy?" *International Journal of Forecasting,* 16(3), 383-397.

1 greater accuracy. In actuality, forecasts based on purchase intentions have
2 sometimes been shown to be more accurate.[22]
3 Why might purchase intention data have yielded the more accurate
4 forecasts? The answer is that while models have their strengths, so do humans. One
5 of the strengths of models is their consistency, yet that is also their weakness when
6 forecasting in a dynamic environment, for "Forecasters can use 'econometric'
7 models effectively only if they have a built-in adjustment mechanism to capture the
8 changing environment."[23] The economy itself was one possible dynamic in the
9 environment at the time that the quantitative data were collected (October 2009),
10 when economic uncertainties were frequent media topics.
11 Another important dynamic for new products is the potential for a
12 competitor to enter the market. Although it is difficult for an econometric model to
13 account in its forecasts for future competitive entry, human judgments are quite
14 flexible in accommodating changing environments.[24] Because as witness Crew
15 noted, market competition is well understood (Tr. VIII/2483, lines 6-8), it is quite
16 likely therefore that a great many participants in the ORC quantitative research did
17 already account for competitive entry, in which case the ORC estimate is one that
18 incorporates possible future competition.

---

[22] Armstrong, J Scott, Vicki G Morwitz, & V Kumar (2000) "Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy?" *International Journal of Forecasting,* 16(3), 383-397.

[23] Hoch, Stephen J. and Robert C. Blattberg (1990) "Database Models and Managerial Intuition: 50% Model + 50% Manager," *Management Science*, 36, 8 (August), p. 887.

[24] Hoch, Stephen J. and Robert C. Blattberg (1990) "Database Models and Managerial Intuition: 50% Model + 50% Manager," *Management Science*, 36, 8 (August), p. 887-899.

So even if an econometric model could be fit to empirical data in order to forecast mail volumes within Five-Day Delivery, it is not at all clear that model forecasts would be any more accurate than those already obtained by ORC. In truth, if an econometric model could have been used, I would suggest that both forecasting methods be used, yielding more information for the present decisions. However, given the lack of empirical data to use with an econometric model, I do not see that such an approach was an available option.

Why a Discrete Choice Model is Not Best

Another analytical approach that has been discussed with the PRC is a discrete choice model (Response to USPS/NALC-T4-17). Discrete choice models have been employed by market researchers to quantify customer preferences for new product features. To gather data for the model, product features are experimentally varied, and research participants select their preferred product profile from a set. For example, an office manager may wish to assess employees' preferences for their company coffee vendor, querying about coffee brands, prices, and locations. Employee respondents would be presented with many choice sets, one example of which might be as shown in Table 3, where respondents would be instructed to select their preferred product. Each column in the table represents a single product.

Table 3: Example of a Discrete Choice Question
Products Shown by Column

Seattle's Best $1.59 / 12 oz Coffee Cart on Your Floor	Starbucks $1.69 / 12 oz Coffee Cart on Your Floor	Starbucks $1.79 / 12 oz Retail Counter in Lobby	Seattle's Best $1.69 / 12 oz Retail Location Across Street

Through a designed experiment that forces participants to make tradeoffs over product features, the analysis reveals their preferences for features such as product brand.

If ORC's objective was to quantify market preferences for Saturday delivery, a discrete choice model would serve well for doing so. However, the objective of the ORC quantitative analysis was to estimate volume, a task for which discrete choice models are not well suited. Discrete choice models directly estimate relative preferences, while volume estimates are absolute quantities. In the coffee example above, respondents are asked which product they prefer, not the quantity they will purchase in a given timeframe.

For volume forecasts using discrete choice models, relative preferences must be converted to absolute quantities, requiring some assumption about the relationship between relative and absolute values. Put differently, a company may be known to have 25 percent market share. To convert market share to sales volumes, additional information is needed.

The additional information might take the form of sales forecasts for one of the products. For example, if an analyst separately estimates a sales volume

forecast for one of the products, the relative measures from a discrete choice model will allow the analyst to estimate sales volumes for the remaining products studied.

In sum, discrete choice models are powerful for understanding relative preferences. Although discrete choice models do not yield sales volumes directly, model results may be used in conjunction with some additional volume estimates for at least one of the products. Such additional volume estimates would need to be estimated with some technique other than a discrete choice model, such as the approach taken by ORC.

Potential Benefit of Additional Research and Analysis is Dwarfed by Cost of Delay

Every analytical approach has its strengths and weaknesses. Since new product forecasts are unable to eliminate the uncertainty of the future, decision makers always have an option to delay a decision and gather additional data and/or conduct additional analyses. As always, the goal of such additional analysis would be to reduce uncertainty further; it can not be eliminated.

So while additional analysis may further reduce uncertainty, additional analysis also comes with a cost, including that of the study itself and another that often looms large in new product contexts: cost of delay until "product launch." So, the decision to conduct additional research or analysis can be framed as a financial decision, the potential additional value of the new information versus the costs of the delay and study.

The benefits and costs of an additional study can be approximated rapidly with ballpark figures. Assume that an additional study would require six months. The cost of delayed implementation of Five-Day Delivery would be half the $3.1B

annual cost savings of the proposal (USPS-T-7, attachment 3, page 2), or $1.55B. The costs of the study itself would increase that figure, so, as a ballpark figure, the costs of conducting an additional study are ">$1.55B."

As for the value of the information, additional study would increase information about revenue reductions from mail volume declines. Currently, volume reduction estimates are estimated to be in the range from 0.35 percent to 1.06 percent,[25] with a mean of 0.71 percent (USPS-T-9, p. 2). Those volume reductions translate into a contribution loss range of about $102M to $308M,[26] with a mean of $206M of contribution loss (USPS-T-7, attachment 3, page 2). As a hypothetical example, an outcome of additional analysis combined with the ORC study may narrow the contribution loss range to: $130M to $292M. Certainly, other ballpark calculations would be plausible as well, for the purpose of ballpark calculations is to assess the magnitude of results. Given this narrowing of a range, which may not even reach a significant digit in the larger context, the costs of additional study of ">$1.55B" outweigh the gain of additional information.

In summary, alternative analytical approaches have been proposed in hindsight. It is not clear whether alternative analyses would be more or less accurate than the ORC study. A critical question at this point in analyzing the Five-

---

[25] This 95% confidence interval was calculated using the standard asymptotic normal formula for percentages, where standard error is square root of pq/n. Here, p is 0.71%, $q = 1-p$, and n is the number of respondents. For n, I used the number of business respondents, 2152. This calculation is used solely as an example for a ballpark estimate. Other standard error estimates could be used in lieu of this one. Of those I considered, all yielded outcomes reasonably close to that given here.

[26] The contribution loss range of $102M to $308M is calculated from the $206M expected loss from USPS-T-7, the 0.71% volume loss from USPS-T-9, and the volume range of 0.35% to 1.06%. Specifically, $102M = $206M*(0.35%)/(0.71%), and $308M = $206M*(1.06%)/(0.71%). These calculations assume a constant marginal contribution for mail volume.

Day Delivery proposal is the value of conducting an additional study. From a financial viewpoint, the costs of an additional study are quite large (>$1.55B), and the financial benefits are, at best, very small relative to the costs. Hence, the answer is quite clear: financially speaking, the Five-Day Delivery proposal should not be delayed for additional estimation of future mail volumes.

I want to note that the estimates cited above are intended to provide a reasonable way to assess the benefits of undertaking additional analysis. They should not be considered as constituting such additional analysis.

CONCLUDING REMARKS

Throughout this testimony, I have maintained a neutral position on the Five-Day Delivery proposal itself. I have focused my analysis on the financial implications alone, knowing that such a decision is not entirely based on finances. If the decision were made on finances alone, I see the clear and compelling answer to be that the USPS will save multiple billions of dollars by adopting Five-Day Delivery. The results of the financial analysis definitively show that the cost savings to the USPS dwarf the revenue reductions.

I do wish to state that I agree with witness Crew (NALC-T4, p. 12) that, in the bigger picture, the Postal Service should consider these steps to be a near-term rather than a long-term solution. In the long term, the Postal Service is best served by a focus on additional ways to add value to customers and other stakeholders such as employees. Not only is the Postal Service valued by its customers (USPS-T-9, p. 4), but it also has expertise and extensive distribution networks in place, giving

1 it a solid basis from which to continue innovating for the benefit of all. I understand

2 that the Postal Service intends to continue seeking ways to add value for its various

3 stakeholders, as outlined in the Postal Service's "Delivering the Future" action plan.