

USPS-SRT-4

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

MAIL PROCESSING NETWORK
RATIONALIZATION SERVICE CHANGES, 2011

Docket No. N2012-1

SURREBUTTAL TESTIMONY OF
REBECCA ELMORE-YALCH
ON BEHALF OF THE
UNITED STATES POSTAL SERVICE

June 22, 2012

Table of Contents

I.	Introduction.....	1
II.	Transit Time as a Critical Aspect of Service	6
III.	Understanding the Concept of Probability.....	10
IV.	Application of a Weight Reflecting the Likelihood that a Change in Behavior Will Occur	16
V.	Inappropriate Calculation of Confidence Intervals.....	24
VI.	Alternative Research Methods	32
VII.	Conclusion	36

1 **I. Introduction**

2 National Association of Letter Carriers (NALC) witness Crew (NALC-T-1)
3 provides testimony in this proceeding that nominally opposes the Postal Service
4 Request in this docket. The Postal Service Request for an advisory opinion regarding
5 the Mail Processing Network Rationalization (MPNR) relies upon testimony that
6 juxtaposes estimated mail volume loss and its consequent impacts upon revenue and
7 contribution, with estimated annual operating savings to show that MPNR provides
8 lasting financial benefits for the Postal Service. Lost mail volume is estimated at 1.7
9 percent—a number that has remained constant throughout the proceeding—while
10 estimates of savings have varied but stayed in the range of low single digit billions of
11 dollars annually. As such, MPNR presents financial gains of known magnitude and sign
12 that justify, at least on financial grounds, its pursuit by the Postal Service.

13 Witness Crew relies upon his opinion and simple economic principles to support
14 his testimony, while refusing to engage in scientific and technical discourse or actual
15 examination and analysis of data, going so far as to say, “I had already found what I
16 considered to be a fundamental flaw, so I didn’t feel it was an appropriate use of my
17 time to get into the analysis of the data and so on.”¹ As such, witness Crew’s testimony
18 does not undermine the Postal Service’s conclusions regarding estimated mail volume
19 loss and its consequent impacts upon revenue and contribution. This surrebuttal
20 testimony focuses only upon witness Crew’s unsubstantiated opinions regarding the
21 market research—an area of expertise he agrees that he has never developed or
22 studied—and its estimates of mail volume losses projected from implementation of

¹ Tr. 11/3618, lines 11-13.

1 MPNR. This surrebuttal testimony accordingly explains why witness Crew's testimony
2 (NALC-T-1, Tr. 11/3542), and his failure to address or even consider the technical merit
3 underlying criticism of his unsupported opinions, make his testimony unhelpful to
4 resolution of the technical issues the Commission's forthcoming advisory opinion will
5 likely address.

6 In the testimony below, I address five sets of issues raised by witness Crew.
7 First, I show that witness Crew's opinion that any change in transit time for some portion
8 of First-Class Mail constitutes a significant decrease in service quality (NALC-T-1, p. 4,
9 footnote) lacks any credible supporting evidence. His opinion was formed without any
10 review of the research report or transcripts from the qualitative research phase (USPS-
11 LR-N2012-1/26). Further, witness Crew provides no support for his opinion in the form
12 of other research, peer-reviewed journal articles or anything of technical merit. His
13 opinion thus cannot be said to be the result of a thoughtful or scientific method.

14 In contrast, the research conducted to support the Request in this docket clearly
15 suggests that transit time *per se* is a relatively unimportant service attribute. I provide
16 representative quotes from the qualitative research conducted for this docket illustrating
17 participants' responses to several questions relating to this issue. I also present results
18 from several studies, including longstanding research about users of the mail, which
19 shows that transit time (or speed) is a less important element of service than reliability,
20 convenience, and cost in determining what type of service postal customers choose to
21 meet their needs.

22 The issue addressed herein is witness Crew's assertion that the concept of
23 probability is not well understood by survey respondents (NALC-T-1, pp. 9 – 10)

1 notwithstanding its widespread use in everyday parlance and in survey research. I
2 present findings from a number of studies illustrating that the majority of survey
3 respondents do understand the concept of probability.

4 The third contention of witness Crew that I address is his continued assertion that
5 it is inappropriate to use a weight to reflect self-reported likelihood of behavior, which
6 implies that a less accurate forecasting method should replace what was used. Witness
7 Crew's support for his opinion that unweighted volume estimates of behavior (volume
8 and product use) should be used is based solely on his prior testimony (NALC-T4,
9 Docket No. N2010-1, pp 5 – 7) and a Commission advisory opinion. Witness Crew
10 does not take into account witness Prof. Peter Boatwright's testimony (USPS-RT-1,
11 PRC Docket No. N2010-1) as well as subsequent support for use in the form of journal
12 articles documenting the applicability and use of a likelihood scale. In my supplemental
13 testimony, I provide common and well-documented examples where use of a likelihood
14 weight demonstrates its superiority over any other method that witness Crew might
15 conceivably have proposed. I further summarize related key findings from numerous
16 peer-reviewed journal articles by well-known and respected academics with
17 documented experience in their respective fields—including market research. As I have
18 testified in this proceeding, the Commission's reluctance to accept the adjustment in
19 Docket No. N2010-1 remains contrary to accepted market research practice.

20 The fourth issue that I address relates to witness Crew's assertions regarding
21 confidence intervals. Witness Crew's opinion that confidence intervals are calculated
22 incorrectly relies upon an economic theory, i.e., the assumption no mailer would
23 increase volumes in light of the proposed change in service. As I will show, this theory

1 is contradicted by real world experience. Crew's contentions also overlook the fact that
2 the majority of respondents indicated that there would be no change in their volume and
3 some reported an increase. Witness Crew offers only his personal opinion to support
4 making an arbitrary judgment in his testimony that all mailers will respond negatively or,
5 under cross-examination, a moderated view that the likely response would be no
6 change in volume or a negative change. Witness Crew indicated in prior testimony that
7 he has essentially no direct experience as a large volume mail shipper, nor has he
8 conducted any research related to how mailers react to service changes.

9 Witness Crew suggests instead that the confidence intervals should have been
10 right-censored but provides no explanation as to how right-censored confidence
11 intervals should be derived and computed, nor does he recognize the implications of
12 that effort. Further, his opinion does not account for additional tests provided by ORC
13 International in response to a request from the Presiding Officer which demonstrate that
14 for half of the estimates, the change in volume is not statistically different than zero.
15 This finding confirms that MPNR will have little impact on overall mail volume and that
16 the majority of businesses and consumers are unlikely to make any significant change
17 in their mailed volumes beyond what is occurring as a result of other changes in the
18 operating environment.

19 The fifth issue that I address relates to witness Crew's reliance on additional
20 measurement tools, the one used in much of the professional work he reviews, such as
21 econometrics or other tools. In the abstract, more information supporting an opinion
22 can reduce the uncertainty inherent in any given decision. However, in reality, the cost
23 of additional information and analysis can far outweigh its benefits. Market research

1 properly designed and conducted by a reputable firm is commonly used as the
2 foundation for real market decisions including the measurement of customer response
3 to a change to an existing product. The Postal Service relies on such market research
4 in this docket.

1 **II. Transit Time as a Critical Aspect of Service**

2 Dr. Crew incorrectly asserts that an increased transit time for First-Class Mail
3 (FCM) will be perceived by customers as a significant decrease in the level of service,
4 and therefore a price increase that will have significant adverse impact on the use of
5 FCM. Contrary to witness Crew’s assertion, the Postal Service has submitted evidence
6 that transit time and arrival speed are not highly salient to customers using FCM apart
7 from the obvious need to predict transit time and be confident that a bill payment will
8 arrive prior to its due date. This predictability, however, rests on a rough knowledge of
9 transit times, such as allowing a week for payments, rather than a sure and certain
10 knowledge of when a payment will actually arrive.

11 To illustrate, FedEx developed a research-based “hierarchy of horrors,” more
12 formally called a service quality index (SQI) in which speed *per se* is not even a factor.
13 A partial proxy for speed, wrong or right day late service failures, garners six points out
14 of a total of 51. Reportedly, reducing SQI scores by 50 percent has helped the
15 company grow volume by 80 percent over the same four-year period,² so it appears to
16 have provided FedEx with real value.

17 In another example from as far back as 1992 (GAO 1993), the Postal Service’s
18 customer satisfaction program, which was audited at Congressional request that year
19 and subsequently approved by the GAO in 1993, focused on factors other than transit
20 time.³ Upon addressing areas of expressed dissatisfaction, the Postal Service was able
21 to improve its aggregate ratings over time. This early report, titled *Tracking Customer*

² Sarah Cook, *Customer Care Excellence: How to Create an Effective Customer Focus*, 5th edition. (New York: Kogan Page, 2008).

³ *Tracking Customer Satisfaction in a Competitive Environment*, General Accounting Office Report. GAO/GGD-93-4.

1 *Satisfaction in a Competitive Environment*, and a history of research since, suggest that
2 predictability and consistency are more important for Postal Service FCM customers
3 than speed *per se* can ever be, given that other market alternatives long ago positioned
4 themselves as the high-speed, more expensive alternatives to the Postal Service. That
5 is, the segment of customers for whom speed is a critical concern has long since relied
6 on shipping and communication methods other than First-Class Mail.

7 Moreover, since email delivery is nearly instantaneous, and Pew (2012) reports
8 that 92 percent of online adults (who comprise 80 percent of the total adult population)
9 use e-mail⁴, clearly a large majority of FCM customers are aware that e-mail is the
10 fastest option for delivery of many items (*e.g.*, documents, letters, notes) that were
11 traditionally sent by First-Class Mail. Moreover, the growing popularity of online bill
12 payment via banks, credit unions and creditors themselves suggests that FCM is a less
13 desirable alternative for those who pay bills just prior to their due date.

14 Thus, it can be reasonably asserted that at this point in the FCM service's life
15 cycle, it is the format of paper that drives residential FCM usage: reactively, if someone
16 receives a paper bill that must be paid or is more easily paid by paper—for example, a
17 one-time payment where one does not want to create an account, or proactively, if
18 someone prefers to communicate her thoughts in a more personal way than e-mail
19 permits—or for a more general example, by sending a greeting card. The same can be
20 said for business usage to the extent that businesses choose to send paper rather than
21 use the option to communicate and transact online.

⁴ Pew Internet Project tracking survey. The survey was fielded from July 25-August 26, 2011. Accessed at <http://pewinternet.org/Reports/2012/Digital-differences.aspx> on June 12, 2012.

1 Research by the American Customer Satisfaction Index (ACSI) in 2011 noted
2 that:

3 Customer satisfaction with the Postal Service's regular mail delivery also
4 improved over last year, up 4 percent to match its former high point of 74. But
5 this gain comes at a time when the volume of mail is shrinking and the Postal
6 Service faces financial difficulties. Indeed, higher satisfaction with the Postal
7 Service might reflect a dwindling customer base, the most loyal of whom are also
8 the most satisfied. The more dissatisfied customers may already have left.⁵

9 This analysis comports with the alternatives available to FCM customers: e-mail or chat
10 programs for "instant gratification" or FedEx/UPS and other package carriers for faster
11 shipments of documents or parcels. If a guarantee of overnight or second day delivery
12 is needed, these carriers gear their entire operations to fulfill that guarantee, as
13 evidenced by FedEx's SQI and the extensive marketing communications that such
14 carriers focus on their speed and reliability.

15 ORC International's qualitative research for the Postal Service is replete with
16 verbatim comments that knowing a piece of mail will arrive essentially the following
17 week is sufficient when sending time driven items such as bill payments, because these
18 are typically processed as they arrive and sent back out to ensure sufficiently timely
19 receipt. Indeed, respondents commented that their current mail transit times do not
20 always seem related to the distance traveled. Moreover, participants explicitly stated
21 that they were more concerned with a continued confidence in ultimate delivery than
22 speed *per se*, just as the other research above suggests. A typical comment:

⁵ The American Customer Satisfaction (ACSI) Commentary, June 2011 issue, accessed at http://www.theacsi.org/index.php?option=com_content&view=article&id=256&Itemid=335. The ACSI is implemented by the American Society for Quality, the University of Michigan and the CFI Group.

1 ***It wouldn't affect me because I do the same thing. I get my bills and***
2 ***usually within a couple of days I sit down. I probably do any bills that***
3 ***come in twice a week. I [sit and] mail it out. I don't wait until the last***
4 ***day. [Chicago, Moderate Income Consumer]⁶***

5 To the extent the concern gets expressed that lowering the service standards
6 would connote loss of dependable arrival, postal communication need simply reassure
7 customers that the element of predictability, the most critical component, still exists. Dr.
8 Crew, a professor of economics, not management or marketing, ignores the impact of
9 appropriate communication of any and all service changes in his testimony. However,
10 service changes in any regulated utility are always “faced out” with appropriate
11 information so that users can take the new standards in context. In fact, his concern
12 that increasing transit times which are already inaccurately perceived as higher than
13 they are “may herald the death knell” for the Postal Service (NALC-T-1, p. 3, line 8)
14 actually supports our conclusion that consumers are not making decisions based on
15 transit times, real or imagined, and the qualitative data from the interviews showing that
16 it would take a significant degradation from the status quo’s transit times to create a
17 perception of inadequate quality, is in accord with that conclusion. More critically, since
18 a more efficient network usually generates greater reliability, it is likely that customer
19 satisfaction will actually increase because the greater reliability is perceived as a service
20 improvement. Dr. Crew’s view of the world, limited as it is to a projection of simple
21 economic theory, never conceived that MPNR heralds a service quality improvement.

⁶ ORC qualitative research, 2011; see USPS-LR-N2012-1/26.

1 **III. Understanding the Concept of Probability**

2 Witness Crew’s statement that “I am not convinced that the concept of probability
3 is well understood by most survey respondents” (NALC-T-1, pp. 9-10; Tr. 11/3551-52) is
4 unsubstantiated. When asked to support this statement with authoritative sources, he
5 stated, “I have not researched this matter so am not aware of any authoritative sources
6 to support my view” and that “[he] does know that the risk associated with various
7 hazards is imperfectly perceived by individuals”⁷ and in support of this statement cites
8 the entirety of a 387-page book to support this statement.

9 The concept of probability is frequently used in survey and market research, from
10 voter surveys to product development, and its uses are well documented. Moreover,
11 research clearly shows that survey respondents understand the concept of probability.
12 By way of example, the National Science Foundation regularly conducts a study that
13 measures public attitudes toward and understanding of science and engineering.⁸ One
14 set of questions specifically measures respondents’ understanding of scientific inquiry
15 scales, of which the understanding of probability is one. To be classified as
16 understanding probability, survey respondent have to answer correctly the following:

17 ***A doctor tells a couple that their genetic makeup means that***
18 ***they've got one in four chances of having a child with an inherited***
19 ***illness.***

20 ***(1) Does this mean that if their first child has the illness, the***
21 ***next three will not have the illness? (No); and***

⁷ Crew Response to USPS/NALC-T1-7(a), Tr. 11/3585.

⁸ National Science Foundation, National Center for Science and Engineering Statistics, Survey of Public Attitudes Toward and Understanding of Science and Technology (1999, 2001); University of Michigan, Survey of Consumer Attitudes (2004); and University of Chicago, National Opinion Research Center, General Social Survey (2006, 2008, 2010).

1 **(2) Does this mean that each of the couple's children will have**
 2 **the same risk of suffering from the illness? (Yes)**

3 Over the years, two out of three respondents answered both questions correctly,
 4 illustrating that the majority of survey respondents do indeed understand the concept of
 5 probability. To illustrate the extent of understanding, if **all** respondents were simply to
 6 guess their responses (thereby suggesting they do not understand probability), the
 7 index would be 25 percent.

	1999 (n=1,882)	2001 (n=1,574)	2004 (n=2,025)	2006 (n=1,864)	2008 (n=2,021)	2010 (n=1,454)
Understanding of probability	64%	67%	64%	69%	64%	66%

8 This research further shows that the understanding of probability measured in
 9 this way increases with education. Given that the majority of those surveyed by ORC
 10 are owners or managers of small, medium, and large businesses and are likely to have
 11 higher education attainments than the general population, it can safely be assumed that
 12 their understanding of probability is higher than the National Science Foundation
 13 numbers.

14 Witness Crew’s own hypothetical example of the likely individual response to a
 15 question regarding the result of a coin toss exercise itself demonstrates that people do
 16 understand the concept of probability.⁹ He supposes that if asked to give an estimate of
 17 how many times out of 100 a coin flip would land on heads, most people would say

⁹ NALC-T-1, p. 9; Tr. 11/3552.

1 50.¹⁰ He then states that if a subsequent question asked people to give an estimate of
2 the likelihood that it the number would actually be 50, most would say something less
3 than 100. If real, this outcome clearly shows that people do understand probability and
4 that the actual occurrence of exactly 50 heads is uncertain. Very few, including many of
5 those well-versed in statistics, would know that the real probability of exactly 50 heads
6 is only 8 percent; the point here is that even Crew's illustrative hypothetical
7 demonstrates that people do understand probability.

8 Further demonstration that most people do understand probability comes from
9 the Health & Retirement study, a source of longitudinal and cross-sectional data for
10 researchers and policymakers who study aging. Conducted periodically by the
11 University of Michigan starting in 1992, this study asks respondents a number of
12 questions about their perceptions concerning chances of future events, such as working
13 past the age of 62 or living to age 75 or 85. Respondents use a scale of 0 to 10, to
14 indicate the chances of such events.¹¹ Over the years, researchers have analyzed this
15 research to assess its validity. Hurd and McGarry (1995) analyzed responses to
16 determine if they behave like probabilities of survival and if their averages are close to
17 the average probabilities in the population.¹² They found that responses to these
18 questions do behave like probabilities and that they do aggregate to population
19 probabilities—that is, when respondents provide a 0 through 10 value on this scale, they
20 are providing a probability based on their experience, knowledge and perceptions.

¹⁰ I doubt that actual replication of this hypothetical would generate this answer, but such digression is not helpful here.

¹¹ See <http://hrsonline.isr.umich.edu/> for additional information regarding the Health & Retirement Study..

¹² Hurd, Michael D. and McGarry, Kathleen, "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study," *The Journal of Human Resources*, Vol. 30, Special Issue on the Health and Retirement Study: Data Quality and Early Results (1995), pp. S268-S292.

1 Moreover, they found that an individual's estimate of the probability they would live to 75
2 or 85 co-varies with other variables in the same way actual outcomes vary with these
3 variables. For example, respondents with higher socioeconomic status reported a
4 higher probability of survival, which is consistent with actual outcomes. Similarly,
5 respondents who smoked reported a lower probability of survival, also consistent with
6 actual outcomes.

7 More recently, Smith, Taylor, and Sloan used four waves of the Health &
8 Retirement Survey to test whether longevity expectations match actual mortality at the
9 individual level.¹³ Like Hurd and MacGarry, these authors found that an individual's
10 subjective estimate of their survival probability is consistent with same individuals'
11 observed survival patterns. They also found that an individual's longevity expectations
12 respond negatively to serious or new health shocks and to improvements in an
13 individual's functional limitations. They conclude that an individual's self-reported
14 longevity expectation is a fairly accurate index of personal survival probability, both in its
15 responsiveness to events that experts would suggest do affect the odds of death and as
16 a prediction of future mortality.

17 Moreover, there is ample research, notably in the healthcare field, which clearly
18 demonstrates that individuals are well aware of and understand probability, especially
19 when expressed as the risk associated with various behaviors. For example, the risks
20 associated with obesity are well-known internationally. A recent study in Mexico, a
21 country with the second highest prevalence in the world of obese adults, examined

¹³ Smith, W. Kerry, Taylor, Donald A., Sloan, Frank A., "Longevity Expectations and Death: Can People Predict Their Own Demise?" *The American Economic Review*, Vol. 91, No. 4 (September 2001), pp. 1126-1134.

1 knowledge concerning obesity co-morbidities. In all groups studied, the researchers
2 found high understanding that obesity is a disease and its causes, together with its
3 consequences such as the greatly increased risk of type 2 diabetes, high blood
4 pressure and knee osteoarthritis.¹⁴ It should be noted that awareness of and capability
5 of assessing risk is so well-documented that the Federal Drug Administration produces
6 a Medication Guide that presents probability information on the risks associated with
7 taking different medications.

8 Finally, Charles F. Manski is both an economist and an expert in survey
9 research. His article in *Econometrica* has been cited hundreds of times in support of
10 using survey respondents' probabilistic expectations to predict behavior. Manski
11 suggests that early research¹⁵ criticizing the use of survey research "predisposed
12 **academic economists** to draw the broad but unsubstantiated conclusion that all data
13 on expectations are suspect." [Emphasis added.]¹⁶ Manski then provides examples of a
14 large number of studies where survey respondents are asked different forms of
15 questions to elicit their expectations or probability of behavior, using some type of
16 probability or likelihood scale, and concludes that "survey respondents are willing and
17 able to report expectations in probabilistic form." Manski properly points out that it is
18 theoretically possible to, rather than ask for a single probability, to ask respondents for a
19 range of probabilities. However, he offers no literature supporting the proposition that
20 asking for a range would increase the precision of the resultant responses. Thus, while

¹⁴ Soriano R, Ponce de León Rosales S, García R, García, E, Méndez JP., "High Knowledge About Obesity And Its Health Risks, With The Exception Of Cancer, Among Mexican Individuals", *Journal of Cancer Education*, 2012 Jun;27(2):306-11

¹⁵ Charles F. Manski, "Measuring Expectations," *Econometrica*, 2004, 1329-1376.

¹⁶ Dr. Crew appears to personify such academic economists based on both his testimony and the statement that "We [including Dr. Kleindorfer] do very little econometrics. We get to read a lot of econometrics as a result of the books we edit. But we do very little econometrics." Tr. 11/3672, lines 5-8.

1 we have no evidence that offering respondents the opportunity to provide a range of
2 probabilities would result in more projectable or usable data, we can state with
3 confidence that academic and scientific literature establishes the reliability and utility of
4 respondent-produced probability estimates. Moreover, since we are asking for a
5 probability, we can and should do so naturalistically in our society, where probabilities
6 are most commonly expressed colloquially as single points, e.g. “There is a 20 percent
7 chance of rain,” “We have a 50 percent chance of getting this proposal accepted,” or
8 “We have a 1 in a million chance of winning the lottery.”

9 Manski directly criticizes economists’ hostility to the use of subjective data and
10 concludes “by and large, persons respond informatively to questions eliciting
11 probabilistic expectations for personally significant events.” He further states, “We have
12 learned enough for me to recommend that economists should abandon their antipathy
13 to measurement of expectations. The unattractive alternative to measurement is to
14 make unsubstantiated assumptions.”

15 To conclude, common sense and this review contradict witness Crew’s claim that
16 individuals in general and survey respondents in particular **fail to understand the**
17 concept of probability.

1 **IV. Application of a Weight Reflecting the Likelihood that a Change**
2 **in Behavior Will Occur**

3 Witness Crew emphatically states that the market research is “flawed” as a result
4 of the “inappropriate” use of weighted volume estimates reflecting respondents’ reported
5 likelihood of changing mailed volume or products used in response to service standard
6 changes. This subject drew significant attention in discussion of the 5-Day Delivery
7 research as well.

8 First, we would like to clarify that multiplying respondents’ stated likelihood of
9 making a change by the estimated change in volume is not an “adjustment.” Rather it is
10 a weight applied to estimates of volume reflecting respondents’ stated likelihood (or
11 probability) that their behavior would change. With businesses, we are in essence
12 asking the likelihood (or probability) that they would change their business operations if
13 new service standards were introduced. Nearly all forecasts of future volumes use
14 some form of weight, using either historical or survey data or in some instances, a best
15 guess.

16 Witness Crew was asked his opinion regarding the common application of this
17 type of weight, based on self-reported likelihood of voting, to forecast an election
18 outcome.¹⁷ He responding by stating that that since results from this type of research
19 are a binary choice, it is different from an **estimate**,¹⁸ suggesting that results from
20 election polls do not provide candidates or the public with an **estimate** of who is likely to
21 win an upcoming election or that results from voter studies do not provide policy-makers

¹⁷ See, e.g., Tr. 11/3585 (regarding interrogatory USPS/NALC-T1-7).

¹⁸ Tr. 11/3645. To help illustrate the paucity of Crew’s reasoning, “estimate” is emboldened in this section.

1 with an **estimate** of the likely success of a forthcoming ballot measure. To the extent
 2 that witness Crew suggests that the results from this type of research or any research
 3 are anything but **estimates**, I disagree. A review of polling methodologies in the U.S.
 4 and U.K. provides clear evidence that weighting an estimate of candidates' potential
 5 election results is widely used and that a failure to do so results in misleading estimates
 6 and conclusions. Polling companies in the UK are required to disclose their
 7 methodologies and are used here for a reference.¹⁹

8 To illustrate the importance of weighting an estimate of behavior, consider the
 9 following simple example of a typical voter survey in which two questions are asked:

- 10 **1. Are you likely to vote for Candidate A or Candidate B?**
 11 **2. What is your likelihood of voting in the next election?**
 12 **Use a scale from 0 to 10 where "0" means "certain not to vote" and**
 13 **"10" means absolutely certain to vote."**

14 Analysis of this data is generally done at the individual level but the following 2x2
 15 table illustrates its application:

	Candidate A	Candidate B
% Would Vote For	60%	40%
% Likely to Vote in Upcoming Election	30%	60%
% of All those Surveyed Voting For	18%	24%
Adjusted Forecast for the Election Outcome	43%	57%

¹⁹ http://news.bbc.co.uk/2/hi/uk_news/politics/vote_2005/basics/4275273.stm.

1 In this example, failure to apply a probability weight reflecting likelihood of voting
2 responses to the **estimate** of the percentage voting for each Candidate would lead one
3 to forecast that Candidate A would win by a significant margin. However, since
4 proponents of Candidate A appear to be less motivated to vote, application of the
5 likelihood of voting weight results in a different and more accurate **estimate** of the likely
6 outcome. Further, by not applying the likelihood of voting weight, Candidate A might
7 use the wrong marketing and communications strategy. Based on the unadjusted
8 results, Candidate A might choose to highlight the differences by “bashing” Candidate
9 B. If voting behavior held true to the reported probabilities of voting, using this approach
10 Candidate A would have to increase the preference margin over Candidate B to 67% /
11 33% in order to win. On the other hand, using a “get out the vote strategy,” Candidate A
12 would only need to increase the percentage of those who favor Candidate A that are
13 likely to vote to 40 percent in order to win.

14 This approach is also used when estimating charitable donations. Potential
15 donors are again asked two questions:

16 **1. How likely are you to donate money to [NAME OF**
17 **CHARITY]?**

18 **2. What is your likely donation amount?**

19 Let's assume that a respondent said in response to Question #1: 25 percent and
20 in response to Question #2: \$100. When projecting to estimate the total likely receipts
21 of the charity we would infer that on average, people like Respondent Y would
22 contribute \$25 each. This will be a far more accurate projection than assuming they all
23 contribute \$100 each. In fact, as in the voting example discussed earlier, failure to

1 apply this weight to the donation estimates could cause a charity to make strategic
2 business decisions that would negatively affect both the charity as well as potential
3 beneficiaries of the charity. Imagine if a charity used its research without applying this
4 weight to determine how to allocate its budget across its beneficiaries. And then imagine
5 if they do not have the money that they expected. Application of this weight is
6 supported by a number of studies on charitable giving and is routinely used. For
7 example, a recent study was conducted to determine the influence of a number of
8 factors including attitudes, norms, perceived behavioral control, and past behavior on
9 intentions to donate money to charitable organizations. Respondents completed a
10 questionnaire assessing these constructs. Four weeks later, a subsample of
11 respondents reported their actual donating behaviors. Results showed that donating
12 intentions were the only significant predictor of donating behavior.²⁰

13 This concept of using a reported probability to improve the utility of an estimate is
14 also applied to economic forecasts, a context with which one might ordinarily expect
15 witness Crew to be more familiar. For example, the 2012 Empire State Manufacturing
16 Survey,²¹ conducted by the Federal Reserve Bank of New York, asks business
17 executives to provide the “percentage chance” that (1) their prices paid and (2) prices
18 charged will increase at one of two levels, stay within 2 percent of current levels or
19 decrease at one of two levels. These data were not used to recalculate potential price
20 changes but were reported as sample averages and are tracked over time, reflecting a
21 perceived need to refine the predicted price changes by a likelihood level.

²⁰ Smith, J. and McSweeney, A., “Charitable Giving: The Effectiveness of a Revised Theory of Planned Behavior Model in Predicting Donating Influences and Behavior,” *Journal of Applied Social Psychology*, 17, 363-386 (2007).

²¹ http://www.newyorkfed.org/survey/empire/empire2012/2012_05Supplmental.pdf.

1 And as already stated, this approach is applied to market forecasts. Studies by
2 Hamilton-Gibbs, Esslemont, and McGuinness (1992)²² and Seymour, Brennan, and
3 Esslemont (1994)²³ provide specific examples where purchase level estimates were
4 weighted by the probability of buying any amount of the product. Respondents in this
5 research were asked to provide estimates of the most likely quantity of six grocery items
6 they would purchase in the next four weeks. They were also asked the purchase
7 probabilities (using variations of the Juster Scale). Predicted purchase levels were
8 then calculated by multiplying the purchase amount by the probability that the
9 respondent will buy that amount. While the overall purpose of these studies was to
10 measure the effects of different types of intention (probability) questions, the results are
11 significant and relevant in that they provide a clear picture of how well the application of
12 the likelihood adjustment approach provides accurate purchase forecasts.
13 Respondents in this research were re-interviewed 28 days after the initial interview and
14 the actual purchase amounts were obtained for each of the six items. Results showed
15 that in several instances, predictive validity was quite high. Differences were generally
16 a function of the product category.

17 Finally, Vicki Morwitz, Leonard N. Stern School of Business at New York
18 University, provides a summary of the extent to which consumers accurately predict
19 behavior.²⁴ She begins by stating that “purchase intentions are routinely used in
20 marketing research to predict whether or not consumers will purchase products.” Using

²² Hamilton-Gibbs, D., Esslemont, D., and McGuinness, D., “Predicting the Demand for Frequently Purchased Items,” *Marketing Bulletin*, 1992, 3, 18 – 23.

²³ Seymour, P, Brennan, M., and Esslemont, D., “Predicting Purchasing Quantities: Further Investigation of the Juster Scale,” *Marketing Bulletin*, 1994, 5, 21-36.

²⁴ Vicki G. Morwitz, “Why Consumers Don’t Always Accurately Predict Their Own Future Behavior,” *Marketing Letters* 8:1 (1997), 57 – 70.

1 a five-point scale to measure intent, she further states that if “purchase intentions were
2 perfect predictors of subsequent purchase, then the conditional probability that
3 consumers engage in a behavior, given that they say they ‘definitely will buy’ will equal
4 one ($p(\text{behavior} \mid \text{intent} = 5) = 1$) and zero for consumers who state they ‘definitely will
5 not buy’ ($p(\text{behavior} \mid \text{intent} = 0) = 0$).” In general, this probability, $p(\text{behavior} \mid \text{intent})$,
6 represents an unbiased measure of intent if the probabilistic measure of intent equals
7 the probability of engaging in a behavior. It is true that in some instances purchase
8 intentions are not unbiased and that in some conditions survey respondents may under-
9 or over-state actual purchase rates. Morwitz examines factors that moderate the
10 accuracy of $p(\text{behavior} \mid \text{intent})$ which thus provides some insights as to when and how
11 this approach should be used.

12 Applicable to this research, Morwitz suggested that consumers having previous
13 experience with a product or service are more accurate predictors of their future
14 behavior than other consumers. The reasoning is that experienced consumers should
15 be better able to assess the pros and cons of engaging in a behavior and to understand
16 factors that will influence their ultimate decisions than inexperienced consumers.

17 To support this prediction, Morwitz and Schmittlein, found that past use of a
18 durable good improved the accuracy of stated future intention. Specifically, they found
19 in regards to stated purchase intentions for a personal computer in the next six months,
20 48 percent of those with previous computer experience at work or school fulfilled their
21 stated intentions compared with only 29 percent of those lacking previous experience.²⁵

²⁵ Morwitz, Vicki and Schmittlein, David, “Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which ‘Intenders’ Actually Buy?,” *Journal of Marketing Research* 29 (1991), 391-405.

1 These results and a review of other studies led them to conclude that “the accuracy of
2 $p(\text{behavior} \mid \text{intent})$ increases with greater experience with the behavior.”

3 Clearly, nearly all consumers and businesses in the United States have at least
4 some experience with FCM. By screening in this research for the person in the
5 household or business with the most knowledge and experience, we can safely state
6 that respondents in this research are the best forecasters of their future responses to
7 changes in postal services. Moreover, as the research discussed above illustrates, the
8 Postal Service and the Commission can have high levels of confidence both that the
9 ORC International research paradigm reflects the highest of survey market research
10 standards and that the results accurately project changes in mailing behavior.

11 In light of this research, a return to witness Crew’s coin tossing hypothetical is
12 constructive. Witness Crew states that an individual’s estimate of how often 100 coin
13 flips would land on heads is 50 followed by an estimate that the probability of actually
14 achieving 50 heads is a hypothetical 80 percent. He states that it would be “obviously
15 wrong to multiply this uncertainty factor of 80 [percent] by 50 to conclude that the
16 respondent’s best estimate of the number of heads would be 40.” While he is
17 superficially correct (because this calculation excludes the potential for more than 50
18 heads), it would be fair to combine the estimate (50) and probability (80%) and project
19 that the individual’s best estimate of the number of times 100 flips of a coin would land
20 on heads would be between 40 and 60 (or that the average number of heads would be
21 50—although this does not make use of the probability estimate).

22 This example is not entirely analogous to ORC International’s research in this
23 docket because the example uses a distribution rather than a single point estimate. In

1 the research, respondents are asked the likelihood of their making a
2 behavioral/operational change and then for a series of estimates as to what their
3 volume would be. We are simply and logically saying that if on average respondents
4 say there is a 50 percent probability they would likely change their behavior or business
5 operations but if they did do so their change in volume would be 100 pieces, it is more
6 accurate to assume that on average the actual change in volume would be 50 pieces
7 rather than 100 pieces. Additionally, we have provided extensive support for this
8 assumption and the application of a weight as confirmed by extensive academic,
9 scientific and peer reviewed sources.

10 In conclusion, virtually every study and model we looked at applies some kind of
11 discount factor on self-reported volume estimates. The research conducted uses a
12 well-documented scale to assess likely behavior. Application of this self-reported
13 behavior, shown to be more reliable than actual behavior, is clearly more appropriate
14 than using the total volume projections without any consideration of what is likely to
15 occur or as what is often done an arbitrary discount figure.

1 **V. Inappropriate Calculation of Confidence Intervals**

2 Witness Crew makes several assertions in his written and oral testimony
3 regarding the calculation of confidence intervals that require clarification and further
4 examination.

5 First, Crew incorrectly asserts that ORC International did **not** calculate
6 confidence intervals for the 5-Day Delivery research and that the only estimate of
7 confidence intervals was provided by Dr. Peter Boatwright in his testimony (USPS-RT-1,
8 pg. 26). These statements are incorrect. In response to a direct request from the Chair
9 during the hearing on July 21, 2010 in the 5-Day Delivery case (Docket No. N2010-1),
10 ORC International subsequently computed confidence intervals for each of the
11 individual estimates of percentage change in volume for each product, which included
12 complete documentation as to how those confidence intervals were computed.²⁶ These
13 results were provided in advance of Dr. Crew's rebuttal testimony on the 5-Day Delivery
14 research so he should have been aware of and familiar with them; he did not then raise
15 his essentially trivial concerns that those confidence intervals included and crossed
16 zero, which is consistent with the fact that no intervenors, including Dr. Crew himself,
17 then raised any such concerns. Dr. Boatwright's testimony simply sought to provide
18 additional insights into impact of the volume loss for the 5-Day Delivery proposal and he
19 noted that other methods for estimating standard error could be used in lieu of his
20 example. Dr. Boatwright does not suggest that this estimate would be the single and
21 sole replacement for the computation of confidence intervals provided by ORC
22 International for this analysis, as witness Crew's testimony supposes. Witness Crew's

²⁶ http://www.prc.gov/Docs/69/69940/Variance_Percent_Change.pdf.

1 statements regarding the previous testimony are incorrect, and suggest that he has not
2 performed an adequate review of documents in the public record to support his opinions
3 and conclusions.

4 In calculating confidence intervals for the current Network Rationalization and
5 First-Class Mail Service Standards research, ORC International used a classical
6 confidence interval calculation detailed in our response to POIR Question #5, Question
7 #24). Witness Crew is correct in that use of this calculation implies there is a chance
8 that the true value could be greater than zero²⁷. However, the appropriate way to
9 address this is subject to some debate.

10 Based on questions raised by witness Crew and the Presiding Officer, ORC
11 International delved more deeply into the subject and located a comprehensive review
12 article on how to compute confidence intervals when there are natural limits to the
13 estimate.²⁸ A summary of this paper and its implications for this research follows.

14 Witness Crew states unequivocally in his testimony that an increase in volume
15 following the proposed change in service is “nonsensical.” (NALC-T-1, p. 13) If one
16 assumes that is true, one possible procedure proposed by Cowen and Ellison would be
17 to treat data points that lay outside the feasible range as unusable, discarding them and
18 then computing a confidence interval. Another approach would be to discard those data
19 points that are outside the feasible range (i.e., greater than 0) and replace them with the
20 nearest feasible value (i.e., 0).

²⁷ NALC-T-1, p. 13

²⁸ Cowen, S. and Ellison, S., “Reporting Measurement Uncertainty and Coverage Intervals Near Natural Limits,” *The Analyst*, May 2006.

1 In this case, such “data censoring” or “shifting” would be inappropriate since it
2 would have a dramatic effect on the apparent estimates and variance. Analysis of the
3 results clearly shows that the majority (64 percent to 89 percent) of consumers and
4 businesses report zero change in their volume and 9 percent to 20 percent report a
5 decrease. However, this analysis also shows that between 0 percent and 18 percent
6 report an increase in their volume, which is consistent with an understanding that the
7 changes proposed would improve the reliability of mail delivery. It is important to note
8 that in the questioning process, respondents were re-read their volume estimate before
9 the proposed change so that their predicted new volume level was anchored by the
10 previous response, thereby focusing attention upon the sign of their response (increase
11 or decrease).

12 The following table clearly illustrates that the vast majority of consumers and
13 businesses report that the proposed change in First-Class Mail service standards will
14 have no impact on the volume of mail they send.

% of Respondents Reporting

National Accounts (n = 26)	Decrease in Volume	11%
	No Change in Volume	89%
	Increase in Volume	0%
Premier Accounts (n = 416)	Decrease in Volume	12%
	No Change in Volume	85%
	Increase in Volume	3%
Preferred Accounts (n = 407)	Decrease in Volume	10%
	No Change in Volume	88%
	Increase in Volume	2%

% of Respondents Reporting

Small Businesses	Decrease in Volume	18%
	No Change in Volume	64%
	Increase in Volume	18%
Home-Based Businesses	Decrease in Volume	18%
	No Change in Volume	66%
	Increase in Volume	16%
Consumers (n = 8670)	Decrease in Volume	22%
	No Change in Volume	78%
	Increase in Volume	<1%

1 As stated above and illustrated in this analysis, witness Crew’s assumption that
2 response to the proposed change in service can only result in a zero or negative volume
3 change is logically and tangibly false. Comments from the qualitative research suggest
4 that the proposed service levels proposed via Network Rationalization for some
5 participants constituted improvement over current expectations for transit time, which
6 would connote a performance improvement, with a potentially synergistic perception
7 that a more efficient network could increase service reliability; as illustrated above,
8 service reliability is by far a more important attribute to mailers than transit time.

9 Therefore, there is every justification for avoiding the censoring or shifting of data
10 that witness Crew proposes, and it would be contrary to standard research practices to
11 do so. Cowen and Ellison agree that neither censoring nor shifting individual

1 observations is appropriate. Instead, the recommendation is to use all data to compute
2 the estimates and standard uncertainty.²⁹

3 Cowen and Ellison then suggest alternatives for computing a confidence interval.
4 The first, simplest, and most widely used is simply to truncate the classical confidence
5 interval. A truncated interval is obtained by first calculating a classical confidence
6 interval using the result and its standard uncertainty and then removing the confidence
7 interval that lies outside the feasible range, effectively shifting it to its natural limit. The
8 other confidence interval is usually well inside the feasible range and is left unchanged.
9 They note that this approach does lead to a modest statistical bias that is increasingly
10 evident as the mean value moves further from zero. In this case the mean estimated
11 value (*i.e.*, percentage change in volume) is very close to zero; therefore, the truncation
12 approach would have little statistical bias and can be used. Using this approach, the
13 confidence intervals provided are correct; one would simply ignore all values greater
14 than zero.

15 This analysis, and as witness Crew correctly points out as well (Tr. 11/3554,
16 NALC-T-1, p. 12), also illustrates that distribution of the estimated percentage change in
17 volume is not normally distributed. Other approaches use methods that are not reliant
18 on distribution assumptions such as bootstrapping to compute a confidence interval.
19 This computationally intensive approach is gradually becoming more feasible as it is
20 incorporated into PC-based statistical applications.

21 As an alternative and in response to a request from the Presiding Officer in this
22 docket (POIR No. 6, Question 3), ORC International undertook additional analysis to

²⁹ *Id.*, p. 712.

1 determine whether the estimates provided were statistically significant from zero. Per
2 this request, ORC International used an approach similar to the Wald test requested by
3 the Presiding Officer. Witness Crew does not anywhere acknowledge that he is aware
4 of or has reviewed this additional analysis.

5 The Postal Service's approach and results are fully documented in our response
6 to POIR No. 6, Question 3.³⁰ In summary, however, ORC International used a test that
7 is comparable to a Wald Test that measures whether the percentage change in volume
8 for each product is statistically significant from zero. As some respondents indicated an
9 increase in volume and others a decrease, it was determined that a two-sided test is
10 appropriate. This analysis found that in more than half (54 percent) of the tests, the
11 results are not statistically significant from zero. While this result does not equate to a
12 conclusion that there will be zero change in volume for these products, we can say that
13 the estimates provided are consistent with the overall conclusion that the impact of the
14 proposed changes is likely to be negative but small.

15 The table below shows the results of this additional analysis. A "fail" indicates
16 that the percentage change in volume for that product is not statistically different from
17 zero; a "pass" means that the percentage change in volume is statistically different from
18 zero.

³⁰ http://www.prc.gov/Docs/81/81816/Answers_POIR6_Qs_3-6.pdf

Account / Customer Segment

	National	Premier	Pre-ferred	Small Business	Home-Based Business	Con-sumers
First-Class Mail	FAIL	FAIL	PASS	FAIL	FAIL	PASS
Presort FCM	FAIL	FAIL	PASS			
Priority Mail	FAIL	FAIL	PASS	FAIL	PASS	PASS
Express Mail	FAIL	FAIL	PASS	PASS	PASS	PASS
Regular Periodical	PASS	FAIL	FAIL			
Not-for-Profit Periodical	(N/A)	PASS	PASS			
Regular Bulk/Standard	(N/A)	FAIL	PASS			
Not-for-Profit Bulk/Standard	(N/A)	FAIL	FAIL			
Total Mail Volume	FAIL	FAIL	PASS	FAIL	PASS	PASS
n/a – Estimate of percentage change was 0% so test is not applicable						

2 Our original approach to computing confidence intervals was to use the classic
3 confidence interval computations used in nearly all survey research, regardless of the
4 underlying distribution of the responses. This approach assumes that the distribution of
5 error around the estimate is normally distributed, which we acknowledge is not correct.
6 However, witness Crew provides no alternatives for the classic confidence interval and
7 does not appear to be aware of other authoritative materials as well as the analyses we
8 have performed. Any suggestion that we would censor or shift the data to address a

1 pre-conceived but unsupported assumption that respondents would only maintain or
2 decrease their volumes as a result of the proposed changes is unwarranted. **Moreover,**
3 **were we to do so, a greater percentage of respondents would report zero change**
4 **in volume, resulting in both a biased estimate and confidence interval.** Instead, it
5 would be appropriate to apply a well-documented truncation of outlying confidence
6 interval values that avoids unwarranted data censoring or shifting. This has no effect on
7 the confidence intervals originally provided by ORC International, except that the
8 upward bound would always be 0%. The additional analysis requested by the Presiding
9 Officer also addresses this concern. By applying this approach, we have supported the
10 result that most respondents did not anticipate a significant change in volume and those
11 that did anticipated a minimal effect.

1 **VI. Alternative Research Methods**

2 Finally, witness Crew makes the same recommendation as he did in 5-Day
3 Delivery that “given that Mr. Whiteman and Ms. Elmore-Yalch are so concerned about
4 bias, it is somewhat ironic that the Postal Service should have also considered
5 alternatives to its survey approach” and suggested “econometric and similar analysis” to
6 predict how a change in service standards would impact mail volume (Tr. 11/3555,
7 NALC-T-1 at 13).

8 First, witness Crew incorrectly suggests that the Postal Service failed to consider
9 alternatives to its survey approach. In fact, the Postal Service and ORC International
10 worked collaboratively throughout the design process to consider alternatives to the
11 survey approach. We incorporated many changes identified in the comprehensive
12 review the 5-Day Delivery research, such as ensuring inclusion of consumers and
13 businesses in Alaska and Hawaii in both the qualitative and quantitative market
14 research, identifying a more “rural” community for the qualitative research than one
15 adjacent to a larger metro area, and splitting the original likelihood question asking for
16 the “likelihood of changing the volume of mail sent or the way in which they send their
17 mail” into two questions to avoid any issues with having a double-barreled question.

18 At the same time, the quantitative research was designed and analysis
19 performed to meet the stated objective to “estimate, by segment, the percentage by
20 which each applicable product’s volume would increase or decrease if changes to
21 current First-Class Mail service standards were implemented.” (USPS-T-11, pg. 15).

22 In his cross-examination, witness Crew demonstrates his lack of knowledge
23 regarding market research and its application by stating:

1 Typically market research is about new product offerings. This isn't really
2 about a new product offering. It's about changing the quality of an existing
3 product. So it is not a common application of market research. (Tr.
4 11/3675, lines14-18).

5 Contrary to this unsupported assertion, market research on existing products is
6 conducted routinely to determine the impact of changes to the product offering, whether
7 it is changing its existing configuration, price or some other attribute. In response to a
8 question during my own cross-examination, I provided an example of research that
9 clearly showed the benefits of market research to predict behavior in response to a
10 decrease in the level of service provided. This article involves a situation where
11 Pennsylvania stocked 28 percent fewer catchable trout in a particular year (due to water
12 quality issues). The analysis showed that survey responses better predicted actual
13 behavior than the econometric analysis. Moreover, the projected response among
14 survey respondents was substantially greater than actual behavior.³¹ Although, this
15 reference was provided well in advance of witness Crew's written testimony and oral
16 cross-examination, witness Crew does not appear to have taken this reference into
17 account.

18 In support of his contention that the Postal Service should have conducted some
19 type of econometric analysis to evaluate change in mailing behavior from service
20 standard changes, witness Crew cites an example of a study combining survey results
21 with econometric studies of demand.³² This study used historical data, including
22 historical data from two panel studies, to model letter traffic demand using time series
23 analysis. This article clearly supports the use and value of econometric models when

³¹ Richard Ready, Donald Epp & Willard Delavan (2005): A Comparison of Revealed, Stated, and Actual Behavior in Response to a Change in Fishing Quality, *Human Dimensions of Wildlife: An International Journal*, 10:1, 39-52

³² http://www.idei.fr/doc/conf/pos/papers_2010/veruete_mckay.pdf

1 historical data are available. However, that is not the case here. To the best of my
2 knowledge, the Postal Service does not have historical data that would enable us to
3 estimate the likely impact of a change in a specific service attribute. If Saturday service
4 had been eliminated as proposed, it could have been possible to use historical data
5 following that hypothetical point to model what a further change in service could mean.
6 But at this point, no comparable service changes exist that could be used in a fashion
7 similar to what Veruete and McKay has done. With respect to doing econometric
8 analysis for the current case, I would also refer to Dr. Boatwright's testimony (USPS-
9 RT-1, pg. 28) where he states that "given the lack of empirical data to use with an
10 econometric model, I do not see how such an approach was an available option."
11 Given Dr. Boatwright's experience and research focus on product development
12 methods, consumer response models, and Bayesian econometric modeling, I believe
13 we can safely assume that the same lack of empirical data to use with an econometric
14 model exists and that therefore such an approach was not an available option for the
15 Postal Service.

16 Finally, witness Crew suggests that some form of experiment in which one group
17 is simply asked for their volume and distribution of volume across products and that
18 group serves as a control group. Then a second group serves as a treatment group
19 and they are presented with the change in service standards and asked what their
20 volume would be. One could then compare the results of the two groups. (Tr. 11/3677-
21 78.) This might be an effective albeit costly approach were the Postal Service actually
22 able to conduct a test of the effect on volume by changing the level of service in

1 selected “test” markets and using a matched set of markets as a control. This approach
2 is simply not feasible in this instance.

3 This approach might also be appropriate if the objective of the study were to test
4 the relative impacts of various strategies such as the potential impact of a single service
5 change—First-Class Mail or elimination of Saturday delivery—or the combined effects
6 of both. For example, respondents could have been randomly assigned to four groups:
7 the control “no change” and three treatment groups (1) change to FCM, (2) elimination
8 of Saturday delivery, and (3) change to FCM and elimination of Saturday delivery.
9 However, as stated earlier the research design of a study is based on the study
10 objectives. The study objectives in this case were to measure the impact of a single
11 change in service on volume and distribution of that volume across Postal Service
12 products.

1 **VII. Conclusion**

2 The Postal Service has a long history of conducting research both to assist in the
3 development of new products and services and to forecast the revenue impact of
4 introducing or eliminating products and services. A substantial amount of attention is
5 paid to ensure that the research conducted meets the specific study objectives and
6 informs key decisions as well as being conducted in a manner that meets high market
7 research industry standards and protocols. The Postal Service has clearly illustrated
8 the depth of thought put into the design, analysis, and reporting of the overall research
9 process. Moreover, although this research effort shows that the change to First-Class
10 Mail service standards may have a negative impact on volume and revenue, the Postal
11 Service has also shown that the cost savings more than offset the revenue losses in a
12 calculus that would continue to show positive returns for the indeterminate future.

13 We do not argue that the volume change will be exactly 1.7 percent overall.
14 However, so far most of the arguments have erred on the side that the volume loss will
15 be greater. It is just as likely, and this research clearly shows, that the volume loss
16 could be less than 1.7 percent and in fact could approach zero because the majority
17 surveyed said they would not change the volume of mail they send or the distribution of
18 this mail across Postal Service products. If so, this would make the cost savings even
19 larger.

20 Finally, it is important to keep in mind that a change in service of this magnitude
21 does not occur in a vacuum. The Postal Service has the opportunity through effective
22 marketing communications to put these changes into a context that makes sense to the
23 customers and minimizes most, if not all, negative impacts. If the network changes

- 1 improve the reliability of First-Class Mail, identified as the most important customer
- 2 attribute, then a well-managed implementation and outreach effort could result in a
- 3 negligible decline in the volume of First-Class Mail.