February 14, 2003

Article 43 – Appendices

Appendix 5 – Mail Survey – Technical Documentation

February 14, 2003

Table of Contents

February 14, 2003

Technical Documentation

I. Overview

In July and August of 2001, a "Hybrid Focus Group/Mailer Survey" was conducted among major U.S. First-Class, Standard A, and Periodicals mailers. The primary purpose of the survey was to gain an in-depth, quantitative understanding of the degree to which these types of mailers would engage in ABA remail in response to changes in the governing UPU provisions. Additionally, there was a desire to identify the extent to which certain liabilities associated with remail (e.g., foreign indicia, limited points of entry, etc.) would curb demand. Finally, information on mailers' awareness of current UPU restrictions, and their perspectives on the effectiveness of these restrictions in preventing remail, was sought.

The survey covered a wide range of topics related to remail potential. Information on current mail volumes, mailing practices, and mailing costs was obtained. Each participant was also asked about their interest in engaging in ABA remail (and the portion of their mail pieces that they would consider sending) under various (hypothetical) implementation scenarios. Additionally, mailers' awareness of remail restrictions was measured, as well as their perceptions of the impact of these restrictions on curbing remail behavior.

Responses from a total of 415 eligible (First-Class, Standard A, or Periodicals) mailers who participated in the survey were used in the analysis. This number includes 24 individuals recruited from focus group sample, who were invited to sessions held in Chicago and New York designed to discuss the nature of the individuals' responses to the survey questions, and to explore in greater depth their reactions to, and reservations about, ABA remail options.

Results of the Hybrid Focus Group/Mailer Survey were documented and included in the Joint Study on Article 43 Final Report, dated May 9, 2002 This Revised Technical Documentation provides more details with respect to the research design and implementation procedures.

II. Survey Sample

- ***Description of the Universe***

  The universe of business locations for the Article 43 Hybrid Focus Group/Mailer Survey consisted of those businesses in the U.S. Postal Service's Corporate Business Customer Information System (CBCIS) file within the continental United States having over 300,000 annual FY2000 mail volume in one or more of the following mail classes: First-Class, Standard A, or Periodicals.

- ***Construction of the Sampling Frame***

  The sampling frame for this study was developed from a CBCIS extract provided by the U.S. Postal Service. The CBCIS is a repository of all United States businesses including current and potential customers, customer profile information, accounting period volume and revenue, and postal product information. The sampling frame was extracted by U.S. Postal Service personnel from the CBCIS, according to the following parameters:

- USPS customers who were permit holders, had meters, and/or were government agency sites; and

- Had positive First-Class Mail, Standard A Mail, and/or Periodicals Mail volume and/or revenue in FY2000 or FY2001.

It was anticipated that these mailing organizations would represent the bulk of the First-Class, Standard A, and Periodicals mail volumes that would be susceptible to remail. During construction of the extract, volumes and revenues were rolled up to a business location or site level such that those with more than one permit for a particular class were aggregated.

The extract included fields for:

a) Business name,
b) Address,
c) Telephone number,
d) DUNS number,
e) SIC codes,
f) Several USPS account and business partner codes, and
g) FY2000 and FY2001 (Year–to-Date) mail volumes and revenues for First-Class, Standard A, and Periodicals mail.

- ***Sample Selection & Sampling Procedures***

Prior to sampling, the frame was originally truncated to include only those businesses with over 500,000 annual FY2000 mail volume in First-Class, Standard A, or Periodicals. *(Note that the 500,000 annual piece threshold applied to each mail class separately.)* This was done to improve sampling efficiency, by concentrating on those businesses where the economic incentives for remail are deemed to be large. During the screening process, however, it was discovered that many locations with CBCIS volumes of 500,000 or greater actually self-reported annual mail volumes below the 500,000 threshold and were, thus, terminated during the screening process. The decision was made, therefore, to expand the sample to include businesses with over 300,000 annual FY2000 mail volumes to ensure that a sufficient number of mailing decision-makers were screened successfully to meet the interview targets described later in this documentation.

Mail volumes associated with locations with over 300,000 annual FY2000 mail volumes account for 97%, 89%, and 84% of total USPS First-Class (excluding single-piece letters and cards), Standard A, and Periodicals Mail, respectively. Individuals who had originally terminated because their location's volume was between 300,000 and 500,000 were recontacted to complete the screening process and to determine if they qualified under the new threshold.

The original CBCIS extract contained 408,713 unique business locations (i.e. sites) with active permits in the appropriate mail classes. After truncating the extract to consider only those locations above the 300,000-piece volume threshold and excluding locations in Alaska, Hawaii, and Puerto Rico, there were 22,953 business locations available to be surveyed. Not all of these locations had valid telephone numbers. Those without valid numbers were sent to Telematch Inc. so that valid phone numbers could be appended to as many records as possible. Overall,

after matching, there were 18,802 records with valid phone numbers available for the survey. Of these, 385 records were set aside for use as sample for the Chicago and New York focus groups. The mailing organizations from this latter group of numbers were screened separately so that available respondents could be invited to a focus group discussion in addition to completing the survey questionnaire on-line. (See Section IV for more details about the focus groups.) The remaining 18,417 records were made available for screening purposes (not all of which were used).

Locations were originally grouped into eleven strata defined by type and volume of mail. Location counts in several of the strata were quite small and these strata were pooled for management of the data collection effort and subsequent weighting activities. A description of the strata and counts of business locations are included in the weighting procedures section of Section V. Prior to data collection, the sample was randomized, and data collection was managed relative to quotas established within the strata to ensure that locations with all combinations of mail classes were represented in the final sample. It should be noted that the preliminary quotas were adjusted during data collection to reflect differential eligibility rates, available sample, and cooperation among eligible respondents.

III. Questionnaire Development & Pretesting

- ***Types & Content of the Survey Documents***

Five types of documents were used for the data collection on this project. These included: 1) screening forms; 2) confirmation e-mail/fax notices; 3) reminder e-mail/fax notices; 4) main survey questionnaire; and 5) focus group discussion guide. These can be found as attachments to this documentation.

- ***Screening Forms:*** The purpose of these forms was to:

  a) Determine if a sampled location was eligible for the study,
  b) Determine if the respondent qualified for interview,
  c) Assign respondents to appropriate sampling cells for recruitment purposes, and
  d) Obtain cooperation and contact information for completing the survey questionnaire.

  The screening form used for the survey alone was constructed to ensure the selection of company locations whose mail volume (not including mail volume sent on behalf of other companies) in at least one mail class (First-Class, Standard A, Periodicals) was 300,000 mail pieces or more. In addition, to be considered eligible, the respondent had to be responsible for a greater percentage of mail volume in at least one of the qualifying mail classes for that company than any other individual at that location. If not, a referral to the individual with responsibility for a greater portion of the mail class was sought.

  The screening form used to recruit focus group respondents made certain that the location sent the required mail volume and also ensured that the respondent was responsible for the majority of the mail class volume being investigated.

- ***Confirmation E-Mail/Fax Notices***: The initial notice was designed to provide respondents with the necessary information for completing the survey questionnaire; that is, the website link, user name, number, and password. Those recruited for the focus groups received a slightly different confirmation letter which not only gave them instructions for the survey, but reminded them of their commitment to participate in the focus group discussion.

- ***Reminder E-mail/Fax Notices***: These notices were designed to alert respondents who had not yet completed the survey to the approaching deadline and encourage them to complete the survey. Also, reminders were developed to encourage those who were in the midst of completing the survey to finish it in its entirety.

- ***Main Survey Questionnaire***: This document was designed to understand current mailing behaviors as well as expected behaviors in response to changes in the UPU provisions governing re-mail. Questions addressed the following issues:

  a) Company characteristics,
  b) Current annual company location mail volumes, as well as mail volumes for which the respondent had responsibility by mail class and mail type,
  c) Use of permits,
  d) Worksharing practices,
  e) Overall and itemized postage and production costs,
  f) Domestic and foreign mail destination areas,
  g) Delivery time ranges,
  h) Use of destination entry discounts,
  i) Volumes of mail sent via USPS' competitors,
  j) Awareness, and perceived effectiveness, of UPU provisions governing ABA remail
  k) Likelihood of engaging in ABA remail for various mail classes and mail types,
  l) Expected time frame for engaging in remail activities, and
  m) Anticipated ABA remail characteristics in terms of delivery time ranges, expected paper and print quality, and likely worksharing practices.

  In addition, the survey used a conjoint exercise (described in detail in Section VI of this document) to assess potential ABA remail volumes under various remail scenarios. Respondents answered questions about different remail scenarios without explicit reference to its legality or illegality. Rather, changes were presented as hypotheticals in order not to bias responses.

- ***Focus Group Discussion Guide***: This document was used by the focus group moderator to probe more deeply into the reasons for specific survey responses and to understand the rationale for mailers' likely ABA remail actions.

- ***Development of the Survey Documents***

  Three distinct activities were undertaken to ensure that the survey documents developed for this project covered the required areas of inquiry and did so in an unambiguous fashion. These included: 1) exploratory depth interviews with key stakeholders; 2) collaborative discussions

with the full project team; and 3) extensive pretesting. The first two activities are addressed here, while the third is described in the next two sections.

Twelve in-depth exploratory interviews were conducted in March and April, 2001, prior to the development of the survey questionnaire. Participants in the interviews included postal wholesalers/consolidators, corporate/institutional mail decision-makers, foreign postal officials, and trade organization representatives, chosen because of their knowledge and influence in the mailing community. Interviews covered a variety of issues related to the current and potential future use of remail. The results of the interviews were used to shape the development of the screener and survey questionnaire.

The development of the screener and survey documents was a collaborative effort between National Analysts and other key project team-members and the sponsoring organizations. Several conference calls and in-person meetings were held in which issues were discussed regarding the desired eligibility requirements for organizations and individual respondents, the content of the survey questionnaire, and the overall structure of the conjoint questions used to explore remail propensities. National Analysts submitted several iterative drafts that were reviewed by the team; revisions were made with each successive draft and final approval was given by the joint team before programming the document.

- ***Administrative Pretesting***

  Prior to programming, the hard copy documents were pretested to ensure that the questions were unambiguous, that the questionnaire flowed smoothly, and that it was not overly burdensome. Several pretests were conducted with First-Class, Standard A, and Periodicals mailing decision-makers from a variety of organizations by National Analysts' project management staff. Each pretest included a thorough debriefing of the respondent, during which the respondent's understanding of individual questions and his/her ability to provide answers was probed extensively. Based on results of the pretests, minor revisions were made to the survey questionnaire, yielding the final version contained in the attachments to this documentation.

- ***Programming & Debugging***

  The screening form was programmed into the Computer Aided Telephone Interviewing (CATI) system while the survey questionnaire was programmed using National Analysts' proprietary NAQuest Internet interviewing software. Both programs were then checked extensively for wording and logic by National Analysts' data processing and project management staff. The programmed questionnaire was then pretested with First-Class, Standard A, and Periodicals mailing decision-makers. These pretests also included an extensive debriefing by National Analysts' project management staff, and further refinements to some of the wording and screen layouts were implemented. The revised, programmed survey questionnaire was then made available to project team members for their final review and comments prior to fielding. The paper version of the "electronic" mail survey questionnaire is attached here.

IV. Data Collection Protocols & Results

- ***Screening Process***

  - ***Screening Overview***

    Screening for the survey was conducted over a period of 32 days from July 9 to August 2, 2001. All survey screening calls were conducted via Computer Assisted Telephone Interviewing (CATI) by experienced telephone interviewers. The screening form was tailored using skip patterns and text replacement commands to insert the appropriate mail class classification based on the respondent's answers to previous questions. Respondents were always asked about relevant mail classes in the following order: First-Class, Standard A, and Periodicals.

    Interviewers asked to speak to "the manager who decides mailing policies and selects or contracts with specific vendors for the organization's primary mail class (such as whether to produce mail pieces in house or outsource the production, what carriers or mail services to use for bulk mailings, etc.)." Interviewers were instructed that possible appropriate titles for such managers included Director/VP Operations, Operations Manager, Controller, Traffic Manager, Production Manager, Distribution Manager, Logistics Manager, and VP of Marketing. According to the screening protocol, referrals to the most appropriate mailing decision-maker were taken, as necessary.

    As an initial screen, company locations were terminated if their primary business was providing mail services for other organizations (i.e., a mail house, messenger service, courier service, or mail delivery service). The rationale for this exclusion was to avoid the double counting of mail volume from those companies who use mail houses, consolidators, and other companies to send mail for them. (Companies who use such mailing support companies were asked to report all their mail volume in the survey, regardless of who actually sent the mail for them.) The remaining locations were screened to exclude any locations where the annual mail volume in each of the three target mail classes (First-Class, Standard A, and Periodicals) was below 300,000 pieces.* Within locations passing this criteria, the individual respondent was considered eligible if they had access to the Internet and were responsible for a greater percentage of mail volume in at least one of the qualifying mail classes for that company than any other individual at that location. Respondents who were eligible decision-makers for two mail classes were assigned to the appropriate multi-class sampling cell, while respondents who were eligible for all three mail classes were assigned to the dual class sampling cell with the lowest recruitment volume at the time of the recruit. Respondents who agreed to participate were informed that they would receive a confirmation letter with instructions for accessing the survey via e-mail or fax within a day.

* As mentioned previously, the original minimum volume threshold was set to 500,000 pieces and, subsequently, adjusted during fielding. Locations screened out under the initial 500,000-piece criteria were then recontacted and invited to participate if their location's volume for a particular mail class met the new minimum 300,000-piece threshold.

– ***CATI Interviewer Training & Monitoring***

An extensive interviewer training and quality control program was employed to assure that accurate data were collected during the screening process. To begin, a data collection team was assigned to the screening portion of the study that included executive interviewers and interviewing supervisors. A training manual was developed for use during training and to serve as reference during data collection (see the attached Training Manual). The interviewers' manual included an overview of the project and specific instructions for administering the screening. Using CATI for the screening interviews and a computerized web survey questionnaire substantially reduced potential data collection errors. As is customary during the CATI screening interview, questions were displayed on a computer screen for the interviewer to read, and the interviewer recorded responses through a keyboard. Use of a computerized screener reduces errors by automatically performing skip patterns and logic tests that are built into the program ahead of time. Interviewers are signaled when responses are inconsistent, so that errors can be corrected with the respondent in real-time.

The screening interviews were monitored closely throughout the data collection period. Each interviewer was monitored by both Data Collection and Project Management staff, and feedback was provided on an ongoing basis. Monitoring was performed via telephone and CRT (computer remote terminal), which allowed the monitor to watch what the interviewer was entering into the computer at the same time he or she was listening to the interview.

– ***Oversight by National Analysts***

National Analysts maintained daily contact with the interviewer supervisors during the screening process, obtaining feedback as to the interpretation of screening questions by respondents, and making small adjustments as necessary to ensure the recruitment of the correct decision-makers.

- ***Invitation Process***

Once respondents agreed to participate in the study, their e-mail address or fax number was recorded, depending on the respondent's desires, and they were sent a confirmation letter within 24 hours. The letter reminded the respondent of the screening phone call and provided them with the following information necessary to complete the survey:

a) Completion deadline,
b) Honorarium amount,
c) Internet address for questionnaire,
d) Instructions for starting the questionnaire program,
e) User name,
f) Password,
g) Contact name,
h) Contact phone number, and

i) Available hours for support services from National Analysts' staff.

- ***Web Survey Completion Process***

  – ***Survey Overview***

  The survey was fielded for only 33 days from July 9 to August 10, 2001 rather than a longer time frame to make certain preliminary results would be available in early Fall, 2001. Respondents were initially assigned to sampling cells according to the screener. Final sampling cell assignment, however, was based on respondents' answers in the mail survey questionnaire.

  The questionnaire was tailored using skip patterns and text replacement commands to fit the appropriate mail class sampling cell classification(s) based on the respondent's answers to prior questions. Questions were displayed on a computer screen for the respondent to read, and the respondent recorded responses using a mouse and keyboard. Use of a computerized questionnaire reduces errors by automatically performing skip patterns and logic tests that are built into the program ahead of time based upon instructions outlined in the paper questionnaire. Respondents can be signaled when responses are inconsistent, so that errors can be corrected. In particular, cases where critical respondent answers in the questionnaire differed from those in the screener, respondents were alerted to this fact and asked to check their current answers for accuracy before continuing with the questionnaire.

  Respondents who were eligible decision-makers for two mail classes were asked questions about both mail classes, while respondents who were eligible for all three mail classes saw questions about only two mail classes chosen at random (i.e., two questionnaire sections chosen from 2-2 [First-Class], 2-3 [Standard A], and 2-4 [Periodicals] of the questionnaire, plus conjoint questions for the corresponding mail classes). The randomization was achieved by taking the respondent ID number and multiplying it by 3, then adding that product to 35 and assigning a negative sign to the resulting sum to generate a seed number. The seed number was then entered into a randomization function created by Visual Basic to generate a random number between 0 and 1. All numbers less than 1/3 were assigned to the First-Class and Standard A mail classes. Numbers between 1/3 and 2/3 were assigned to First-Class and Periodicals mail classes. Numbers greater than 2/3 were assigned to Standard A and Periodicals mail classes.

  Respondents were given a toll-free telephone number to call if they had any questions while completing the questionnaire. These questions were fielded by one of two National Analysts Operations or Project Managers who were intimately involved in the questionnaire development, data collection, or analytic phases of the study.

  – ***Reminders***

  Reminder contacts were initiated beginning July 13, 2001. Respondents who had begun, but not yet completed, the survey were sent a (e-mail or fax) letter (see Section III) reminding them of the importance of completing the survey, offering a bonus honorarium for completion by the deadline, and giving instructions for how to log back on to the

website to complete the survey. Respondents who had not yet begun the survey within several weeks of recruitment were also sent a (e-mail or fax) letter (see Section III) reminding them of the survey, offering a bonus honorarium for completion by the deadline, and repeating the original instructions for completing the survey. All of these follow-up activities were necessary, given the short time period established for data collection.

– ***Supplemental Focus Groups***

During the fielding of the survey (July, 2001), five focus group discussions were conducted in Chicago and New York, as shown in the chart below.

Table 1
Distribution of Hybrid Mailer Survey Focus Groups

Mail Type	New York	Chicago	Total
First-Class Mail	X	X	2
Standard A Mail	X	X	2
Periodicals Mail	X	--	1
Total	3	2	5

All sample records with zip codes near the focus group facilities (385 records) were extracted from the sample file and called for screening the recruitment purposes for the focus groups. From the 385 records, 33 were recruited for the focus groups. From these, data from 24 were used (23 of them having participated in the discussions). The remainder were either terminated or voided because of insufficient mail volume (n=5) or partial questionnaire completion (n=4).

Those who participated in the groups represented a mix of operational and business characteristics, including services provided, non-profit status, use of outsourcing for mail production and preparation, mail type and volume sent, and geographic scope of mailing activities.

As noted previously, each participant was required to complete the mailer survey. Discussions during the groups centered on reactions to the concept of legalized ABA remail implementation, the importance of key factors on their willingness to engage in remail, their awareness and assessment of the UPU restrictions that are currently in place, and the impact that removing these restrictions would have on the Postal Service's business. The survey questionnaires completed by the focus group respondents were treated in the same manner as those of other respondents.

- ***Field Results***

Each day, the results of the CATI screening interviews were downloaded, and the completed questionnaires that had been received that day from the CATI and focus group respondents were uploaded into the questionnaire database. Progress reports were prepared daily to ensure that the sample was being worked according to established protocols and to monitor progress toward reaching the study quotas. The reports included the number of eligible and ineligible

respondents, as well as completed interviews, delineated by location, employee size, and industry type.

At the conclusion of the data collection, a final disposition report was issued, summarizing the results of all attempts and contacts, as shown in Table 2 below.

Table 2
Final Result of Call Codes – Screenings/Interviews

Dispositions	Summary	Final Disposition
Total # Available	18,417	
Not attempted		2,301
Total # Attempted (at least once)	16,116	
Callback scheduled/in progress		5,408
No answers, busy, left message		3,485
Refusals		2,416
Not business working numbers (fax, residences, disconnects)		546
Non-working numbers		511
Not available during field period		247
Language problems		33
Started screening		3,470
Total # Location Screener Started/Not Completed	1,017	
Wrong business		203
Undetermined eligibility		814
Total # Location Screener Completed	2,453	
Ineligible location		1,007
Eligible location		1,446
Total # Attempted to Find Eligible Respondent for Web Survey	1,446	
Referral in progress		116
Refused		84
No Internet access		27
Invited to take survey		1,219
Total # Invited to Take Web Survey	1,219	
Refused		89
Agreed		1,130
Total # Did Not Access Web Survey At All	437	
Total # Partially Completed Web Survey	260	
Ineligible (insufficient volume)		105
Stopped before/during Section II		108
Stopped during Section III		47
Total # Completed Web Survey*	458	
Voided interviews		43
Interviews used in the analysis		415

It should be noted that of the numbers still being worked, an average of 2.7 attempts had been made to contact the company. These calls were made during different dayparts and on different days of the week.

During the fielding of the study, the progress of respondents through the survey was monitored through daily reports which summarized the number of respondents accessing/completing the

* Includes 433 from CATI screening plus 25 individuals recruited from the focus group list. (Note: One questionnaire was subsequently voided, leaving 24 in the analysis.)

survey, the number of completed surveys in each sampling cell, the number of terminated respondents and the reason for termination, the number of incomplete surveys and the screen on which the respondent had stopped, the elapsed time the survey had been accessed, and other key statistics.

As shown in Table 2 above, 458 individuals completed the survey, from which 415 were used in the analysis. Even though 1,130 individuals consented initially, 437 of them did not access the survey during the allotted time period. In addition, depending upon an individual's mail class responsibilities, he/she may have responded for up to two mail classes during the interview. Our 415 usable interviews yielded a total of 534 mail class responses – 199 First-Class, 244 Standard A, and 91 Periodicals – as described more fully in the weighting section following.

V. Data Preparation & Weighting

- ***Editing & Cleaning Procedures***

As is customary, a portion of each screening interviewer's work was monitored. In addition, once collected, the screening and survey data were subjected to a rigorous set of electronic and manual checks. The screening data were run through an electronic cleaning program which verified that the skip patterns and consistency checks built into the CATI screening were working appropriately and that the Result of Call codes (ROCs) that had been assigned to each respondent matched the results of their screening questions, ensuring that only eligible respondents had been recruited for the survey. These basic cleaning checks provided assurance that the CATI program was working correctly, the data layout provided from the interviewing facility was accurate, and that no corruption of the data occurred during the downloading process.

The survey data were also run through electronic cleaning programs which verified that the skip patterns and consistency checks built into the web survey questionnaire were working appropriately. Additional logic checks were also used to ensure that the reported responses were internally consistent and reasonable. The data were checked for unweighted outliers or unusual responses (e.g., much higher volumes than in the CBCIS file, etc.). To that end, separate frequency distributions were produced for every quantitative variable in the questionnaire. These distributions were analyzed and an outlier boundary was designated for each variable at either three standard deviations outside the mean or some other more restrictive large value (e.g., 100,000 pieces of current remail volume, etc.). All respondents whose answers fell very close to or outside the established boundaries were flagged for further inspection. Subsequently, a complete set of responses and flags for each respondent needing inspection and/or a callback was produced. These data were then examined manually on a respondent-by-respondent basis by looking at the printout of all responses; callbacks were attempted by National Analysts' project management staff for each respondent whose answers could not be resolved by a simple examination of their data. During the callback, the staff member attempted to resolve all issues. Callbacks were completed with approximately 150 respondents and corrections made, as necessary. During this process, we determined that 43 respondents' location volumes had been overstated, and were actually less than 300,000 threshold and, as such, these interviews were voided.

In short, these cleaning activities were designed to ensure that the tabulated survey data, particularly mail volumes, production and postage costs, and estimated current and future remail volumes, were accurately reported.

- *Weighting Procedures*

Four sets of survey weights were developed for this survey. A ***location weight*** was developed for each respondent, projecting the sample to the population of eligible businesses with mail volumes above the eligible threshold. This weight was used for analyzing survey questions that were asked of all respondents. Three ***mail-class specific weights*** were also developed (one for First-Class, Standard A, and Periodicals). These weights were used for analyzing survey questions that were asked concerning specific mail classes. Mail class specific weights were required because: 1) in locations with three eligible mail classes above the 300,000 threshold, two mail classes were subsampled to reduce respondent burden; and 2) at some locations, different individuals had responsibility for separate mail classes, and the respondent could not reliably respond to some of the detailed mail-class specific questions in the questionnaire. The class-specific weights were the primary weights that were used in the analysis.

Table 3 presents total frame counts within nine strata defined by the distribution of mail type and volumes on the CBCIS file. Secondary stratification by volume within First-Class Only, Standard A Only, Periodicals Only, and First-Class/Standard A was included in the design to provide better representation of mail volumes within these strata. Stratum boundaries were chosen using the cumulative square root of frequency rule (Cochrane, pp. 128-132), with a stratum-level measure of size reflecting the sum of eligible stratum-level mail volumes above the truncation threshold.

Table 3
Number of Business Locations

Stratum	# Frame Locations
Low First-Class Only	2,586
High First-Class Only	586
Low Standard A Only	11,297
High Standard A Only	2,268
Low Periodicals Only	2,121
High Periodicals Only	500
Low First-Class/Standard A	2,641
High First-Class/Standard A	584
First-Class/Periodicals, Standard A/Periodicals, & First-Class/Standard A/Periodicals	432
Total	**22,953**

– ***Location Weights***

An initial set of base location weights was produced where the weights were equal for respondents within strata, and that projected respondents up to the estimated population of eligible business locations. These weights were subsequently raked so that self-reported volumes for each mail class matched volume control totals from the CBCIS file. Base location weights can be viewed as being created from four components: 1) the inverse of the sampling fraction in the stratum, which is the ratio of the number of business locations with valid telephone numbers in the stratum divided by the number of locations in the sample that were actually released and worked during data collection; 2) an adjustment for non-response at the screener level, calculated as the ratio of the number of released and worked sample records, divided by the number of completed screener interviews in each stratum; 3) an adjustment for non-response at the questionnaire level, calculated as the ratio of completed screener interviews with locations that were identified as eligible divided by the number of completed interviews within each stratum; and 4) a coverage adjustment, calculated as the ratio of the number of business locations on the sampling frame divided by the number of business locations that were eligible to be included. This final adjustment essentially projects the survey sample up to the estimated population of eligible locations on the entire (truncated) CBCIS file, rather than simply that portion of the file with phone numbers. The base location weight is the product of these four factors.

When evaluating the construction of the base weights, it is important to note that the major weight components are the adjustments for non-response and coverage. The levels of non-response at the screener and questionnaire level are high, despite considerable efforts during data collection to complete interviews with respondents, primarily because of a very restricted time frame for data collection and the need to work so much of the sample concurrently to ensure all required interviews were completed. As a result, it is important to evaluate assumptions concerning these adjustments, to ensure that they properly protect against biases from non-response and non-coverage.

It was not possible given time and budget restrictions to complete a non-response follow-up survey to investigate fully all potential sources of bias and possible corrections. In this situation, non-response adjustments must essentially use information from the sampling frame to make adjustments. While it would be possible to use screener information for the non-response adjustment at the questionnaire level, the screener also collected information concerning mail volumes, and this information was not available for locations on the frame that did not complete the screener. As a result, evaluation of the non-response adjustments was confined to the information on the CBCIS file. To evaluate the screener non-response adjustment, we compared median total FY2000 mail volumes within strata from the CBCIS file for those business locations that were sampled, but did not complete a screening interview, with those locations with identified eligibility status (either eligible or ineligible) as defined in the screening questionnaire. To evaluate the questionnaire non-response adjustment, we compared median total FY2000 mail volumes from the CBCIS file within strata for those business locations that completed the survey with those who were identified as eligible locations, but did not complete the survey. In each case, while there were clearly substantial differences in mail volumes across the strata, medians were similar

for respondents and non-respondents within strata. We also compared median total mail volumes within strata for those locations without phone numbers and those with phone numbers. Again, these estimates were broadly similar within strata.

As a final adjustment, the base weights were raked so that the weighted total reported mail volumes for each mail class matched control totals from the truncated CBCIS file. During the raking process, the final location weights of respondents were constrained so that no respondent represented more than approximately 10 percent of the weighted volume in each mail class. The raking algorithm was a variant of the linear truncated or restricted modified chi-square procedure.* Since the self-reported volume estimates differed for each individual, the raking algorithm defines the raking weights implicitly, and there is no closed form solution or equation that can be used to explicitly calculate the final weights.

– ***Mail Class Specific Weights***

Three mail class specific weights were constructed. These weights projected the mail class specific survey samples up to the population of eligible locations with volumes above 300,000 in the specific mail class. The steps in the calculation of the mail class specific weights were similar to those for the location weights, with two exceptions. First, an additional adjustment was added for subsampling by mail type for those locations that had volumes above 300,000 for all three mail classes. Second, the base mail class weights were raked to match the estimated population of eligible locations with volumes above the threshold within strata, as well as the volume within the mail class, as defined by the final location weights.* This last step insured that the various weights were internally consistent.

More specifically, mail class specific weights were developed from five components: 1) the inverse of the sampling fraction in the stratum, which is the ratio of the number of business locations with valid telephone numbers in the stratum divided by the number of locations in the sample that were actually released and worked during data collection; 2) for those locations that were eligible to be interviewed in all three mail categories, an adjustment factor of (3/2) was applied to correct for subsampling of mail classes; 3) an adjustment for non-response at the screener level, calculated as the ratio of the number of released and worked sample records, divided by the number of completed screener interviews in each stratum; 4) an adjustment for non-response at the questionnaire level, calculated as the ratio of completed screener interviews with locations that were identified as eligible <u>in the mail class</u> divided by the number of completed mail-class specific interviews within each stratum; and 5) a coverage adjustment, calculated as the ratio of the number of business locations on the sampling frame divided by the number of business locations that were eligible. The base mail class specific weights were calculated as the product of these five weighting adjustments.

---

* See A.C. Singh and C. A. Mohl, (1996) Understanding calibration estimators in survey sampling, Survey Methodology, Vol. 22, pp.107-115.

* For First–Class, one additional set of raking targets was included. Information concerning the method of payment for the entire CBCIS was obtained from the USPS. The questionnaire provided detail on three possible methods of payment: 1) own permit, 2) third party permit, and 3) other. Information concerning the percentage of First-Class own/third party (45%) versus other (55%) forms on the entire CBCIS file was also used as a target in the raking procedure.

Table 4 summarizes the distribution of the four sets of final mail class specific weights.

Table 4
Summary Statistics, Location & Mail Class Specific Weights

Weight	n	Min	Mean	Median	Max
Location	415	1.00	33.55	28.44	283.01
First-Class	199	1.00	35.93	21.58	463.95
Standard A	244	1.00	38.32	29.02	288.78
Periodicals	91	1.00	65.46	43.03	689.81

VI. Conjoint Design & Model Development

- ***Rationale for Use of Conjoint***

As outlined in the Statement of Work (SOW), one objective of the research was to determine the impact that a wide range of considerations associated with ABA remail would have on potential remail volumes. In addition to understanding the impact of lower postage and/or production costs, the SOW required an assessment of the following ABA remail risks that domestic mailers would consider:

- Foreign indicia
- Poorer comparative service levels
- The availability of domestic dropship for domestic mail
- The limited number of entry points for mail into the U.S.
- Quality control
- Address correction and undeliverable mail handling

Conjoint analysis is a widely used market research technique that is ideally suited for assessing the relative importance of various product and service features, including price, that affect usage decisions. The different characteristics that accompany a product or service are presented to a respondent in combination, and questions are asked to elicit information regarding the impact of the specified combination of characteristics on respondent behavior. In this particular study, a full-profile, share allocation conjoint approach was used to elicit information regarding the amount of ABA remail that individual respondents would send, if any, as a function of the characteristics of the remail service that mailers would receive. In this manner, the importance (i.e., effect on remail volume) of mailing cost savings, time-in-stream, and several other factors (including those listed above) was ascertained.

- ***Selection of Attributes & Levels***

The conjoint design employed in this study displayed two ABA remail service scenarios on each screen, each defined using a combination of characteristics describing the nature of the remail service (i.e., remail attributes and levels). Respondents were asked to indicate the amount of their mail they would send via remail under each of the remail service scenarios listed.

The specific set of attributes and levels used to describe the characteristics of the remail options presented was established by the project team. Information assembled by the project team members regarding potential postage and production cost savings, as well as insights gleaned from the individual depth interviews, was used to develop the list of attributes, and determine the most relevant levels to test. A complete list of the remail attributes and levels used in the conjoint analysis is provided in Table 5.

Table 5
Remail Attributes and Levels for Conjoint Analysis

A.	**Brand of Remail Carrier**
1)	Foreign posts or affiliates (e.g., Deutsche-Post, Aero-Mail, Global Mail, etc.)
2)	Branded non-USPS carrier (e.g., FedEx, Airborne, etc.)
3)	Wholesaler/consolidator (e.g., Save-On-Mail, etc.
4)	USPS Domestic (fixed as first option of each screen)
B.	**Official Country of Mailing Origin**
1)	Industrialized European country, such as France or Belgium
2)	Developing European country, such as Hungary, Poland, Czech Republic
3)	Mexico
4)	Canada
5)	Developing Caribbean, or Central/Latin American country
6)	Developing Asian/Pacific-rim country such as Hong Kong, Indonesia, or Thailand
7)	United States (fixed as first option of each screen)
C.	**Indicia/Return Address**
1)	U.S. indicia and return address
2)	Generic international indicia (no specific country name); U.S. return address
3)	Indicia bearing the name of "B" country; return address in U.S.
4)	Indicia bearing the name of "B" country; return address in "B" country
D.	**Mean Time-in-Stream**
1)	2 day average
2)	4 day average
3)	8 day average
4)	14 day average
5)	22 day average
E.	**Variability of Time-in-Stream**
1)	± 1 day
2)	± 3 days
3)	± 5 days
4)	± 10 days
F.	**Address Correction/Forwarding/Mail Handling**
1)	Address correction available, mail forwarded, undeliverable mail returned
2)	No address correction available, mail forwarded, undeliverable mail returned
3)	No address correction available, mail forwarded, undeliverable mail not returned
4)	No address correction available, mail not forwarded, undeliverable mail not returned
G.	**Mail Production Location**
1)	Facilities that your company owns in the U.S.
2)	Facilities in the "B" country that your company owns or would need to purchase/build
3)	Facilities in the U.S. that are operated by a third party producer
4)	Facilities in the "B" country that are operated by an international company
5)	Facilities in the "B" country that are operated by a local independent producer

(Continued)

Table 5
Attributes and Levels

H.	**Mail Preparation Location**
1)	Facilities that your company owns in the U.S.
2)	Facilities in the "B" country that your company already owns or would need to purchase/ build
3)	Facilities in the U.S. that are operated by a third party preparer (i.e. presort house etc…)
4)	Facilities in the "B" country that are operated by an international company (e.g., foreign post, wholesaler, etc.)
5)	Facilities in the "B" country that are operated by a local independent preparer
I.	**Postage Expense**
1)	50% savings vs. current postage expense
2)	20% savings vs. current postage expense
3)	10% savings vs. current postage expense
4)	Current postage expense
J.	**Mail Production/Preparation/Transportation Expense**
1)	50% savings vs. current expense
2)	20% savings vs. current expense
3)	Current printing, production, and transportation expense
4)	20% increase vs. current expense

- ***Description of Design***

  In traditional conjoint analysis, a series of experimentally controlled combinations of attribute levels for a product or set of products are presented to respondents for evaluation (ratings or rankings). A multiple regression model (at the individual respondent level) is then used to model ratings as a function of the experimental design variables. In almost all studies, the number of different combinations of attributes and levels is very large, and it would be impossible (and extremely burdensome) to obtain ratings from each individual regarding all possible combinations. A conjoint design selects a subset of the overall number of possible combinations of attribute levels to be evaluated by the respondent, and the conjoint model allows predictions to be made concerning ratings for combinations of attributes and levels that are not explicitly included in the design.

  As described previously, the conjoint design employed in this study displayed multiple service scenarios on each screen, and respondents were asked to allocate their mail volume within a specific mail class across these options. The first option presented on each screen (Option 1) was merely a description of the current USPS service, as illustrated in the questionnaire document. The second and third options on each screen were ABA remail options, defined by a combination of the attributes and levels described in Table 5.

  When multiple product or service combinations are displayed at one time for evaluation by respondents, the experimental design is a choice set design. The use of choice sets in conjoint studies was first suggested by Louviere and Woodworth (1983).* There are a variety of manual

* Louviere, Jordan J. and George Woodworth (1983) "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, **20** (November) pp. 350-67.

and computer-assisted approaches for generating experimental designs for choice sets.* For this study, a computer-assisted design approach was employed, utilizing the experimental design software resident in SAS procedures FACTEX and OPTEX, and choice set design macros developed in Kuhfeld (2000).**

During the survey, respondents were asked to provide conjoint ratings for at most two mail classes. Our experience suggested that an upper bound on the number of conjoint choice sets that an individual can reliably evaluate during electronic data collection is approximately ten.

To provide coverage of the design space, a design with 5 blocks of 10 choice sets was developed.*** For each respondent, a design was chosen from one of the 5 blocks. Within each block, the order of the choice sets presented in the interview was rotated across respondents to prevent order biases. In situations where the respondent was eligible for more than one mail class, 5 choice sets from each of the two mail classes were shown, for a total of 10 choice sets. In situations where the respondent was eligible for both flats/packages and cards/letters, the respondent allocated flats volume and cards volume separately across each option of the choice set.

The five design blocks are shown in the following tables, where the attributes appear as column headers (with letters corresponding to attributes in Table 5) and active levels of the attributes are specified in rows (with numbers corresponding to attribute levels in Table 5). The option number refers to the columns on the computer screen. Each respondent was exposed to three options at a time. The first was always the USPS, and options 2 and 3 had the attributes and levels shown in the following tables.

* These different approaches are discussed in Chrzan, K., and B. Orme (2000) "An Overview and Comparison of Design Strategies for Choice-Based Conjoint Analysis," 2000 Sawtooth Software Conference Proceedings. A more detailed description of the design approach adopted here is provided in Kuhfeld, Warren, Randal D. Tobias and Mark Garratt (1995) "Efficient Experimental Designs with Marketing Research Applications," *Journal of Marketing Research* **31** (November), 545-57.

** Kuhfeld, Warren F. (2000) Marketing Research Methods in the SAS System, Version 8 Edition, SAS Institute.

*** The overall number of choice sets was chosen to ensure adequate design efficiency (i.e., reliability of estimates), as measured by conventional D-efficiency criteria for linear designs. Grouping of the choice sets is preferable to selecting choice sets randomly from the entire design space because it allows for manual adjustment of the computer-generated designs for level balance. Without this step, an individual respondent might see sets that were biased towards specific levels for certain attributes, and the resulting preference estimates would be skewed.

Table 6a
Conjoint Design Block #1

Choice Set #	Option #	Attributes									
		A	B	C	D	E	F	G	H	I	J
1	2	1	5	2	4	4	4	3	1	2	4
	3	2	1	3	3	2	3	1	3	3	3
2	2	2	2	1	4	2	2	2	2	2	2
	3	3	1	4	4	3	1	4	2	1	3
3	2	1	1	1	1	4	2	5	2	4	2
	3	2	6	4	3	4	4	4	4	1	4
4	2	1	6	3	3	3	1	1	3	1	4
	3	3	4	4	5	2	2	5	2	4	2
5	2	2	5	4	5	4	3	1	5	4	1
	3	3	3	2	2	4	4	3	3	3	4
6	2	1	2	2	1	2	3	2	4	4	1
	3	2	4	2	1	1	1	5	2	2	2
7	2	1	3	4	4	1	3	4	5	3	2
	3	2	5	3	1	3	2	5	4	1	1
8	2	1	3	3	5	1	1	2	2	3	4
	3	3	5	1	3	1	4	2	5	2	1
9	2	1	4	1	2	2	2	3	5	1	3
	3	3	2	3	5	3	1	4	4	2	2
10	2	2	3	1	2	3	4	3	1	4	3
	3	3	6	2	2	1	3	1	1	3	4

Table 6b
Conjoint Design Block #2

Choice Set #	Option #	Attributes									
		A	B	C	D	E	F	G	H	I	J
1	2	3	4	1	3	1	2	5	4	2	2
	3	2	6	4	2	1	4	1	3	2	4
2	2	1	5	2	1	1	1	2	2	3	2
	3	2	4	1	5	4	1	5	4	3	1
3	2	3	6	2	5	3	1	4	5	2	2
	3	2	2	3	1	3	4	5	2	3	3
4	2	2	5	4	3	4	3	4	4	4	1
	3	1	2	1	4	4	4	1	1	1	3
5	2	1	1	1	3	2	2	4	2	3	4
	3	3	2	2	1	4	4	2	2	4	1
6	2	2	3	2	3	3	2	1	5	4	4
	3	3	5	3	2	2	4	3	3	1	3
7	2	2	1	3	1	2	1	2	5	2	2
	3	3	3	1	4	3	3	1	1	1	4
8	2	1	6	3	4	1	3	3	3	1	3
	3	3	1	4	5	1	2	3	1	3	4
9	2	1	3	4	2	3	2	5	4	2	1
	3	3	5	3	2	4	1	1	2	4	1
10	2	2	1	2	4	2	3	3	1	1	3
	3	1	4	4	5	2	3	3	5	4	4

Table 6c
Conjoint Design Block #3

Choice Set #	Option #	Attributes									
		A	B	C	D	E	F	G	H	I	J
1	2	3	6	2	1	2	2	5	5	3	1
	3	1	2	3	4	2	2	3	1	1	3
2	2	3	1	3	4	3	3	2	4	3	2
	3	1	3	4	3	4	1	3	5	2	1
3	2	2	1	2	1	4	1	4	5	1	2
	3	3	2	4	2	2	4	5	2	2	3
4	2	3	2	3	3	1	1	3	2	4	2
	3	2	4	1	2	4	3	3	1	2	4
5	2	1	5	1	2	3	2	1	3	3	4
	3	2	3	2	5	1	3	5	2	2	3
6	2	1	1	2	2	1	3	1	2	4	2
	3	3	5	1	1	3	2	4	5	4	3
7	2	2	6	3	4	1	4	3	3	4	3
	3	3	4	4	4	4	4	3	1	2	4
8	2	1	4	4	5	3	4	2	4	4	2
	3	3	3	1	5	4	1	1	3	1	4
9	2	2	5	3	3	2	2	2	5	1	1
	3	1	6	1	1	1	1	5	4	3	1
10	2	1	4	2	5	2	4	4	2	1	3
	3	2	2	4	3	3	2	1	4	3	4

Table 6d
Conjoint Design Block #4

Choice Set #	Option #	Attributes									
		A	B	C	D	E	F	G	H	I	J
1	2	2	2	1	1	2	4	2	4	4	1
	3	1	3	3	2	4	3	1	1	2	3
2	2	3	6	1	5	4	3	3	3	4	4
	3	2	4	4	2	3	1	1	3	1	3
3	2	3	4	3	1	1	1	5	4	4	1
	3	2	6	1	5	2	2	2	2	2	2
4	2	3	3	2	3	3	4	1	4	1	3
	3	1	6	4	2	2	4	5	5	4	4
5	2	1	1	2	4	1	1	2	4	2	4
	3	3	5	2	4	2	4	3	2	1	2
6	2	1	2	4	3	1	2	3	3	4	3
	3	2	3	3	4	3	3	3	1	2	4
7	2	2	5	2	5	1	4	4	5	1	2
	3	3	1	1	2	1	2	3	1	3	4
8	2	2	1	4	3	4	1	3	2	3	4
	3	1	4	2	1	4	1	4	5	3	2
9	2	3	2	4	1	3	3	2	5	2	1
	3	1	6	3	3	2	2	1	1	3	3
10	2	3	4	3	4	4	2	1	5	3	2
	3	1	5	1	5	3	3	5	2	1	1

Table 6e
Conjoint Design Block #5

Choice Set #	Option #	Attributes									
		A	B	C	D	E	F	G	H	I	J
1	2	1	3	3	1	2	2	4	2	2	4
	3	3	5	1	5	2	1	5	4	4	3
2	2	3	4	1	1	4	3	2	5	3	2
	3	1	1	4	5	3	4	1	3	2	3
3	2	2	6	1	3	3	3	3	3	3	4
	3	3	1	4	4	1	4	1	2	1	1
4	2	2	1	3	1	3	4	2	5	4	4
	3	1	2	2	2	4	2	5	5	1	1
5	2	2	4	2	3	2	4	4	2	1	1
	3	1	4	1	2	1	3	3	5	3	2
6	2	2	3	4	4	4	2	1	3	2	3
	3	3	3	2	5	1	3	3	1	4	4
7	2	3	6	2	4	3	2	3	3	2	4
	3	1	6	4	3	4	3	2	4	1	3
8	2	2	6	2	2	2	1	1	1	1	3
	3	1	1	3	3	1	4	5	4	3	2
9	2	3	2	3	2	3	1	4	2	2	4
	3	2	5	4	1	1	1	2	4	4	2
10	2	1	5	1	4	2	1	1	1	4	3
	3	2	2	3	5	4	2	5	2	3	1

During construction of the designs several prohibitions were imposed that prevented illogical combinations of levels from being displayed. More specifically:

- Mail Production Location levels in the "B" country (levels 2, 4 or 5) were prohibited from being displayed with Mail Preparation Location levels in the U.S.

- Mail Production Location levels in the US (levels 1 or 3) were prohibited from being displayed in combination with a Mean Time-In-Stream level of 2 days

- If Mail Production Location and Mail Preparation Location were both set at levels in the US (levels 1 or 3), then Mail Production/Preparation/Transportation Expense levels were only allowed to take "current" and "20% increase" values

Additionally, it should be noted that while Table 5 presents attributes and levels as main (or marginal) effects, the design was constructed to support the estimation of an interaction (i.e. cross-effect) between mean time and variability of time-in-stream. For low mean time-in-stream, the effect of higher variability is asymmetric (i.e., with a mean of 2 days, ± 5 days was interpreted as providing an anticipated delivery window of 1-7 days), suggesting that an interaction might be appropriate. Both the simpler main effects model and the more complicated model with interactions were estimated and compared.

- ***Preference Parameters***

  Estimates of preference parameters were constructed from the conjoint data using a Hierarchical Bayesian (HB) modeling approach. The HB approach for modeling preference parameters in conjoint research has become very popular in recent years. This popularity has stemmed primarily from evidence that predictions from models estimated using this approach are more accurate than predictions from models estimated using other approaches. In addition, recent advances in Bayesian computing involving Markov Chain Monte Carlo (MCMC) methods have allowed for the application of these techniques for problems of reasonable size (i.e. numbers of attributes and levels). Prior to these developments, Bayesian methods were only really computationally feasible for small problems.

  Under the HB approach, preference parameters are distributed randomly in the population with a common mean. Estimates of the posterior means of the preference parameters or partworths are constructed using Markov Chain Monte Carlo (MCMC) methods. National Analysts has developed a suite of proprietary software applications to generate estimates of partworths from conjoint data using HB modeling techniques. The software routines are written in GAUSS, and are derived from MCMC routines originally developed by Peter Lenk and Greg Allenby.

  The power in the Bayesian approach stems directly from the hierarchical model. In a traditional conjoint model, the number of preference parameters to estimate is large, often approaching the number of conjoint ratings per individual. In this situation, traditional regression-based estimates are very imprecise. The HB estimates, on the other hand, are a mixture of the mean over all respondents and the individual classical regression estimate, with the mixing proportion reflecting the relative precision of the individual classical regression estimates, the sample mean over all respondents, and the difference between the individual estimate and the sample mean. Imprecise individual-level parameter (classical) regression estimates are "shrunk" towards an estimate of the population mean in the Bayesian approach. In tests of out-of-sample predictive performance, the Bayesian estimates typically dominate other approaches by a wide margin.*

- ***Model Development***

  The preference model that was estimated had a tiered structure. Based on input from the qualitative interviews (focus groups and individual in-depth interviews), attributes deemed to be of primary economic importance (net cost, indicia, mean time-in-stream and variability in time-in-stream) determined the base level of remail (i.e., "tier one" attributes). The remaining secondary ("tier two") attributes modified the base level of remail obtained by the first level. Modeling using this hierarchical structure allowed us to impose logical constraints concerning the economic profitability of remail on the demand forecasts. More specifically, remail demands estimated from the first stage of the model were constrained to be zero for any situation where there was no net cost advantage to remail, if the mean time-in-stream was equivalent to or slower than perceived current USPS performance levels. In the absence of this constraint, traditional unconstrained conjoint models could generate illogical results. This section of the

* Lenk, P. J., DeSarbo, W. S., Green P. E. and Young, M. R. (1996) "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15, 173-191.

report provides details on the construction of the preference parameter estimates using conjoint choice task responses. Post-estimation adjustments to the preference parameter estimates and calculation of the remail volumes and share estimates are discussed in the next section.

The model is based on a latent utility specification for the alternatives shown to respondents in each choice exercise. Denote Y_{ijm} as the utility for alternative m on the j[th] choice task, for i=1,…n respondents in the mail class in question. For First-Class and Standard A mail classes, models were developed separately for "cards and letters" and "flats and packages." Respondents were included in the modeling exercise for each component if they were eligible for the conjoint based on the selection algorithm in the questionnaire, and they had positive volumes for the component (i.e., "cards and letters" and/or "flats and packages"), and they allocated a positive amount of remail in at least one scenario. Respondents saw j=1… J_i choice exercises, where J_i =5 if they were selected in two mail classes, and J_i =10 if they were eligible in one mail class. There were m=1,…,3 alternatives shown on each conjoint screen. The preference model assumes that

$$Y_{ijm} = x'_{ijm}\beta_i(\exp(z'_{ijm}\delta_i)) + \varepsilon_{ijm}$$

where x_{ijm} are dummy variables (and an intercept) representing the displayed levels for the tier one attributes associated with each remail supplier profile in the choice set, and z_{ijm} are dummy variables (and an intercept) representing the displayed levels for the tier two attributes. We assume that $E(\exp(z'_{ijm}\delta_i))=1$ so that second tier attributes have no effect on the mean of the latent utility on average, for each respondent. However, variation in tier 2 variables can modify the distribution of the latent utilities around the mean utility. Additionally, the model assumes that $\varepsilon_{ijm} \sim N(0,\sigma^2)$, $B_i \sim N(B,V_B)$, $\delta_i \sim N(\Delta,V_\Delta)$, and that these variables are jointly independent. The model for latent utility is completed with specification of inverse Wishart priors on the parameters $V_B \sim IW(f_0,G_0^{-1})$ and $V_\Delta \sim IW(f_1,G_1^{-1})$

Although the latent utility is not observed, it is possible to compute the parameter estimates using a link function that maps the latent utilities to the reported remail shares in the conjoint exercise. We assume that reported remail shares Y^*_{ijm} are related to the latent utilities in the following manner:

$$Y^*_{ijm} = \frac{Y_{ijm}}{\sum_m Y_{ijm} | (Y_m > 0)} \text{ if } Y_{ijm} > 0, \text{ 0 otherwise}$$

Prior to developing the estimates of preference parameters, three specific cleaning steps were employed to ensure that conjoint response data were appropriate for estimation. First, there were a number of respondents who never selected remail under any scenario, instead allocating 100 percent of their mail volume to the USPS in each case. These individuals were excluded from the conjoint estimation data file; and in the model simulations the brand partworth for these individuals was set to ensure 100% USPS volume allocation under any scenario. Second,

examination of response patterns and choice task timing information by screen that is collected as a diagnostic during interviewing suggested that 6 respondents had ignored the exercises and rated all remail options equally, regardless of their characteristics. These individuals were excluded from the conjoint estimation data file as well, and in the model simulations the partworths for these individuals were imputed using the average partworth for their mail type (First-Class Flats, First-Class Cards, Standard A Flats, Standard A Cards, Periodicals). Third, on a relatively small number of screens (less than 4%), some respondents allocated positive remail amounts to options where their net mailing cost was higher than current, and where the mean time to deliver was not faster than their current service with the USPS. In contrast, results of the qualitative interviews conducted among mailers during this project (focus group and individual depth interviews) indicate that remail would not be pursued if neither net cost nor time-in-stream benefits were available. Accordingly, for modeling purposes, the remail volume associated with these options was treated as a respondent input error for these screens and set to zero, with the remaining allocations rescaled to sum to 100%.

The model was estimated using a modified Gibbs sampler with data augmentation, a Markov Chain Monte Carlo (MCMC) method that generates a sample from the posterior distribution of the parameters. Each set of parameter estimates was estimated using 20,000 iterations of the Markov chain, with the last 10,000 used to estimate the parameters. Convergence was ascertained through starting the chain at multiple start points and ensuring that the chains converged to a common set of parameter estimates, and through the evaluation of time series diagnostics.

With respect to included explanatory variables, only main effects were included in the final model. (During the analysis, interactions between mean time-in-stream and variability in time-in-stream were initially reviewed, but their inclusion did not prove to enhance the predictive performance of the model.) In addition, a "net cost" variable was used that combined postage and mail preparation/pricing/production expenses rather than the separate main effects. In the actual conjoint exercise, respondents were presented with actual levels for postage and mail preparation/pricing/production expenses. In addition, a weighted average net cost change was constructed, where the weights were derived from the self-reported proportions of the two respective cost elements (production/preparation vs. postage) in the respondents' total costs. A categorical representation of the net cost variable was then constructed with four levels (35-50% net cost reduction, 10-35% reduction, 0-10% reduction, and "increased net cost"); midpoints of these ranges were used to define the four levels of this net cost variable in the simulator (see below) that was developed, and serve as the basis for interpolation or extrapolation of net cost reduction values in the simulation analysis.

During model development, the tiered model was compared with a standard linear model where all attributes and levels were included as tier one main effects. However, the standard linear model produced parameter estimates that were counterintuitive. In addition, the standard linear model did not fit as well as the tiered model, when comparing out-of-sample R-squared estimates.* As a result, the tiered model was selected for use in the simulator. We also evaluated a hierarchical model where the population hyperparameters were allowed to vary by

* For out-of-sample goodness-of-fit tests, a hold-out choice set was randomly selected for each respondent, the model was estimated using the remaining choice sets, and the estimated parameters were used to predict the volume shares across alternatives on the hold-out card.

strata. Out-of-sample predictions for this model were also inferior, primarily because sample sizes within strata were smaller after removing constant raters (i.e. individuals with no remail activity under any scenario), effectively reducing the benefits of HB shrinkage on forecast accuracy.

- ***Market Simulator***

The conjoint models provide predictions at the individual level of the share of mail (by mail type) that an individual respondent would allocate to remail, given specified levels of USPS service performance and the characteristics of the remail options. An estimate of the total volume of mail allocated to each remail option is then constructed using the respondent's total volume by mail type, and the final mail-class-specific analysis weight. For each mail type, a spreadsheet simulator was developed to streamline these calculations.

Estimates of remail share are calculated at the individual level using the posterior means of the parameter estimates, the equation for the latent utilities, and the share allocation link, i.e.

$$\hat{Y}_{ijm} = x'_{ijm}\hat{\beta}_i(\exp(z'_{ijm}\hat{\delta}_i)),$$

$$\hat{Y}^*_{ijm} = \frac{\hat{Y}_{ijm}}{\sum_m \hat{Y}_{ijm}|(Y_{ijm}>0)} \text{ if } \hat{Y}_{ijm}>0, \text{ 0 otherwise}$$

These estimates are "plug-in" or "approximate" Bayesian estimates in that shares are calculated as nonlinear functions of the posterior means of the parameter estimates, rather than as an average of the values of the function over the MCMC iterations. Calculating the share estimates as averages over the MCMC iterations would have required that the MCMC algorithm be embedded in the simulator. This would have greatly increased the complexity of the simulation algorithm, and would also have made the simulator run-time prohibitive. The plug-in or approximate Bayesian estimates are typically used for simulators in market research.*

Several post-estimation calibration adjustments were made to the share calculations in the simulator in order to align the model more closely with the scenario characteristics that would ultimately be used as input. Within each simulator, remail scenarios were provided as input for 36 separate options representing electronic and surface mail options in 6 different country groups and 3 branded transportation options. As a result, there were 37 (= 2 x 6 x 3 + 1) mail options overall, when the USPS was included. Since the number of remail options displayed in the conjoint exercise was much smaller than this, an adjustment factor was required to ensure that individual remail shares were not distributed across all 36 remail options. For the share calculations in each simulator, an individual's share of remail was calculated using the most attractive remail option for that individual as the only remail option, i.e.

$$\hat{Y}_{ij} = \max_m(x'_{ijm}\hat{\beta}_i(\exp(z'_{ijm}\hat{\delta}_i))$$

* See Orme, B. and G. Baker (2000) Comparing Hierarchical Bayes draws and randomized first choice draws for conjoint simulators, Sawtooth Software.

for non-USPS remail options and

$$Y_{ijUSPS} = (x'_{ijUSPS}\hat{\beta}_i(\exp(z'_{ijUSPS}\hat{\delta}_i))$$

for the USPS. Estimated shares were then obtained using the share allocation link with only these two options considered. Since each respondent only saw a limited number of remail options in each choice set, this adjustment ensured that estimated shares in the simulator more closely matched stated shares from the conjoint exercise. When aggregated across all respondents, this approach also allocated remail across a larger number of relatively attractive remail options.

A second adjustment was made to the preference parameter estimates to ensure that remail demands estimated by the simulator were constrained to be zero for any situation where there was no net cost or time-in-stream advantage to remail (see above). This adjustment was calculated as the estimated value of utility that would make remail utility zero when net costs and time-in stream were comparable to USPS levels. Since adjusting preferences in this fashion would reduce the attractiveness of remail at all net costs (not just near zero), the constant adjustment was gradually added back (over the range of 0% to 5% net cost reduction) using a piecewise linear function. At net cost reductions greater than 5%, the estimated remail share reflects the original preference estimates with no adjustment. Mathematically, with A_{ij} denoting the adjustment,

$$Y_{ijm} = (x'_{ijm}\hat{\beta}_i + A_{ij})(\exp(z'_{ijm}\hat{\delta}_i))$$

when the net cost advantage to remail as perceived by the respondent is zero or negative,

$$Y_{ijm} = (x'_{ijm}\hat{\beta}_i + A_{ij}(1\text{-}nc_{ijm}/0.05)(\exp(z'_{ijm}\hat{\delta}_i))$$

when the net cost advantage (nc_{ijm}) is between zero and five percent, and

$$Y_{ijm} = x'_{ijm}\hat{\beta}_i(\exp(z'_{ijm}\hat{\delta}_i))$$

otherwise.

Finally, an upper bound was placed on the second tier multiplier to constrain the effect of the tier two attributes during simulations. This ensured that error in the individual level parameter estimates in the second tier did not generate large swings in the estimated utilities and remail shares because of the exponential term. More precisely, the contribution of the second term in the estimate utilities was replaced by

$$\overline{C} \text{ for } \exp(z'_{ijm}\hat{\delta}_i) > \overline{C}$$

After experimentation with the simulators, $\overline{C}$ was set at 2 for each simulator.

It should be noted that the purpose of the Mailer Survey Summary Report was merely to provide an overview of the key information gathered in the survey. As part of this overview, estimates were generated to illustrate the possible extent of ABA remail activity that might occur over a range of potential cost savings (presented in Figure B and Table 10 in the Report), using the models that had been developed from the conjoint data (as described in Section VI). ***However, as identified in the report, these estimates are very preliminary.*** More precise estimates of remail volume were derived from the Remail Impact Model (RIM), which produces aggregate estimates using the conjoint models in combination with detailed information on mail flows and the levels of arbitrage that could exist for various countries and mail classes.

VII. Survey Results & Variance Estimation

- ***Types of Survey Estimates***

The Mailer Survey Summary Report dated December 17, 2001 (see Appendix 4) includes survey estimates for totals, means, percentages, medians, and ratios. When these estimates are for the entire population, the statistics are computed using all individuals in the sample, and the sums in the formulae run from 1 to n respondents. When these estimates focus on a subpopulation of interest, only those respondents in the subpopulation are included in the estimates, and the sums in the formulae run from 1 to the number of respondents in the subpopulation of interest. For example, numbers presented in "Total" columns include values from all respondents. Numbers presented in the "First-Class" column include only values from respondents with a non-zero First-Class weight. "Standard A" columns include only respondents with a non-zero Standard A weight, and "Periodicals" columns include only respondents with a non-zero Periodicals weight. Columns with other headings can be read as follows: rows that correspond to mail-class-specific survey questions used the mail-class-specific weight, rows that correspond to general non-mail class specific survey questions used the location weight.

- ***Formulae (Analytical Expressions) for Survey Estimates***

This documentation denotes W_i as the analysis weight, effectively suppressing the distinction between the location and mail-class specific weights in this section, because the formulae will be identical for each set of weights.

Denote Y_i as the observed value of an analysis variable for the i^{th} respondent. For a categorical variable C with l levels ($c_1, c_2, \ldots, c_l$), denote Y_i^k as an indicator function with $Y_i^k = 1$ if the observed value of variable $C = c_k$, 0 otherwise.

The estimated mean or average value of an analysis variable Y is computed as

$$\bar{Y} = \sum_{i=1}^{n} W_i Y_i / \sum_{i=1}^{n} W_i$$

For estimated proportions, the same formula is used with the indicator function Y_i^k replacing the analysis variable Y, i.e.

$$\hat{P}=\sum_{i=1}^{n} W_i Y_i^k / \sum_{i=1}^{n} W_i$$

Estimated totals are calculated as the weighted sum over the sample,

$$\hat{Y}=\sum_{i=1}^{n} W_i Y_i$$

The report also includes estimates of ratios, where an analysis variable X is divided by analysis variable Y. For example, estimated remail volume shares are calculated as the estimated total volume of remail for the mail class in the scenario under investigation (estimated using the conjoint modeling described in Section VI), divided by the estimated total volume for the mail class. Estimated ratios are calculated as:

$$\hat{R}=(\hat{X}/\hat{Y})=(\sum_{i=1}^{n} W_i X_i) / (\sum_{i=1}^{n} W_i Y_i)$$

Additionally, the report includes estimates of the median number of company locations, the median number of employees at the current location, and the median number of total employees in the organization (Tables 3 and 8 in the Hybrid Focus Group/Mailer Survey – Mailer Survey Summary Report). The actual survey questions were categorical in nature, so an assumption was made that there was a uniform distribution across the discrete integer values within the categorical groupings, and interpolation was used to report individual median values.

Finally, in Table 10 of the Hybrid Focus Group/Mailer Survey – Mailer Survey Summary Report, estimates of the impact of remail considerations on remail volumes are presented. The values presented are (the negative of) the absolute value of the difference in estimated aggregate remail volume share (based on the models described in Section VI) between: a) baseline scenario (defined by an estimated 15.9% remail share, assuming a 25% net cost savings, 6-day (±3 day) time-in-stream from completed production, foreign indicia, either U.S. or foreign production, and no address correction service); and b) alternate scenarios constructed to isolate the effect of each remail risk listed in Table 10.* The specific adjustments to the baseline scenario used to determine the differences presented in Table 10 are as follows:

- *Foreign Indicia*: Changed remail indicia characteristics to U.S. Indicia and return address
- *Poorer Comparative Service Levels*: Improved remail mean time-in-stream and variability in time-in-stream to perceived current USPS levels
- *Domestic Dropship Availability*: Increased remail net cost saving on Standard A and Periodicals to 30% for all respondents using dropship on 50% or more of their mail pieces
- *Limited Entry Points*: Decreased remail time-in-stream to 5.7 (±.2.3) days (based on Data Tabulations, December 17,2001, Table 14)

* Aggregate remail volume share is the volume-weighted share of current USPS First-Class, Standard A, and Periodicals volume [excluding First-Class single-piece mail] projected to be sent via remail using the Conjoint model described in Section VI.

- *Quality Control*: Changed remail to require foreign production (i.e., no U.S. production)

- *Address Correction & Undeliverable Mail*: Changed remail to provide address correction

It should be noted that the purpose of the Mailer Survey Summary Report was merely to provide an overview of the key information gathered in the survey. As part of this overview, the market simulator produced from the survey was used to illustrate the possible extent of ABA remail activity that might occur over a range of potential cost savings (presented in Figure B and Table 10 in the Report). The results presented in that report used the models that had been developed from the conjoint data (as described in Section VI). ***However, as identified in the report, these estimates were very preliminary and did not reflect all of the additional insights and computations/adjustments from the Remail Impact Model (RIM).*** The more precise estimates of remail volume, derived from the Remail Impact Model (RIM), use detailed information on mail flows and the levels of arbitrage that could exist for various countries and mail classes in combination with the conjoint results.

- ***Variance Estimation Procedures***

Respondents in this survey have unequal weights because of the sample design and the various weight adjustments outlined in the weighting section. Variance calculations from standard software that are appropriate for simple random samples will therefore provide incorrect variance estimates when applied to these data. Special purpose software such as SUDAAN, STATA, WESVAR, or PROC SURVEYMEANS in SAS must be used to properly calculate variance estimates for statistics of interest in this survey. Variance estimates in the report were computed using PROC SURVEYMEANS in SAS.

Variance calculations for this survey are complicated somewhat further because survey weights were raked so that weighted mail volumes matched control totals. Variance estimation becomes even more complicated when raking has been employed. The variance estimates in this report use the raked weights, but otherwise ignore the effect of raking in the variance calculations. This approach allows the use of convenient Taylor-series linearization variance estimators that are available in standard software. However, the variance estimates presented here will be conservative (i.e. too large), relative to the variance estimates that would arise if the raking procedure were completely reflected in the calculations.

It is also important to note that the variance calculations above refer exclusively to sampling error. As described in Section VI, a conjoint exercise was also included in the survey to allow for the modeling of responses to hypothetical remail scenarios, and the predictions from the model are subject to modeling error for which there is also additional uncertainty. No attempt has been made to quantify this source of uncertainty and combine it with the sampling variability in the report. Since the conjoint models were estimated from a Bayesian perspective, inferences concerning remail volumes under different hypothetical scenarios could be based on the posterior distribution of these estimates computed using the assumptions underlying the model. Under this alternative approach, the inferences would combine both sampling and modeling uncertainty, although the assumptions underlying the computations would be different than those used in the calculation of the variances for the survey responses. However, under this

alternative approach, it would still not be possible to reflect the post-estimation adjustments to the model discussed in the model development section. Perhaps more important, inferences under this alternative approach would require separate Gibbs sample runs for each scenario, greatly increasing the complexity of the simulator and making run time prohibitive. As a result, variance estimates for estimated remail volumes under different scenarios were not computed from the posterior distribution using the Gibbs sampler.

Measures of sampling error for remail shares included in the report were calculated using the predicted values from the model, and conventional formulae for variance estimates from sample surveys. Traditional within-sample R-squared measures for the models range between 80 and 90 percent for all respondents across mail types, indicating that the conjoint model predicts individual remail choices fairly accurately within sample. From this perspective, the classical measure of sampling variability applied to predict values from the conjoint model should only slightly understate the uncertainty underlying the remail estimates for any scenario.

Finally, it is also important to remember that there may be additional non-sampling sources of error (e.g., over- or understatement of remail intentions, reporting omissions, etc.) that contribute to uncertainty regarding the remail estimates, and those are not reflected in the estimates of sampling error.

- ***Variance Estimation Formulae***

Estimated variances for means, totals, and ratios in the report are calculated using formulae for domain means, totals and ratios as included in PROC SURVEYMEANS in SAS.

For the variance estimates, it is useful to add an h subscript to the variable identifiers indicating that the respondent belongs in stratum h, i.e. Y_{hi} is the response value for the *i*th respondent in stratum *h*, W_{hi} is the respondent's analysis weight. Let n_h be the number of respondents in stratum h, and define I^d_{hi} to be the indicator function for domain d, i.e. I^d_{hi} =1 if observation belongs to domain d, and zero otherwise. Let $Z_{hi}=Y_{hi}I^d_{hi}$, and $V_{hi}=W_{hi}I^d_{hi}$. Estimates of domain means, totals, and ratios can be computed as:*

Means:

$$\bar{Y}^d=\sum_{h=1}^{H}\sum_{i=1}^{n_h}V_{ih}Z_{ih}/\sum_{h=1}^{H}\sum_{i=1}^{n_h}V_{ih},$$

*Strictly speaking, these formulae assume that sampling is with replacement, which is reasonable only if sampling fractions within strata are relatively small. The estimates also condition on n_h in the target population, rather than the overall population that was screened for eligibility. Regarding n_h, this approach is often undertaken in surveys that screen to obtain interviews from a target population that is a fraction of the universe originally sampled for screening.

Totals:

$$\hat{Y}^{d}=\sum_{h=1}^{H}\sum_{i=1}^{n_h} V_{ih} Z_{ih},$$

Ratios:

$$\hat{R}^{d}=(\hat{X}^{d}/\hat{Y}^{d})=\sum_{h=1}^{H}\sum_{i=1}^{n_h} V_{ih} Z_{ih}^{x}/\sum_{h=1}^{H}\sum_{i=1}^{n_h} V_{ih} Z_{ih}^{y}$$

The associated variance estimators are:

Means:

$$\hat{V}(\bar{Y})=\sum_{h=1}^{H}\frac{n_h}{n_h-1}\sum_{i=1}^{n_h}(\tau_{hi}-\tau_h)^2$$

where

$$\tau_{hi}=(V_{hi}Z_{hi}-\bar{Y}^{d})/\sum_{h=1}^{H}\sum_{i=1}^{n_h} V_{ih}$$

$$\tau_h=(\sum_{i=1}^{n_h}\tau_{hi})/n_h$$

Totals:

$$\hat{V}(\hat{Y})=\sum_{h=1}^{H}\frac{n_h}{n_h-1}\sum_{i=1}^{n_h}(Z_{hi}-Z_h)^2$$

where

$$Z_h=(\sum_{i=1}^{n_h} Z_{hi})/n_h$$

Ratios:

$$\hat{V}(\hat{R})=(1/\hat{Y})^2[\hat{V}(\hat{X})+\hat{R}^2\hat{V}(\hat{Y})-2\hat{R}\hat{C}ov(\hat{X},\hat{Y})]$$

where

$$\hat{C}OV(\hat{X},\hat{Y})=\sum_{h=1}^{H}\frac{n_h}{n_h-1}\sum_{i=1}^{n_h}(Z_{hi}^{X}-Z_h^{X})(Z_{hi}^{Y}-Z_h^{Y})$$