University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Travel and Tourism Research Association: Advancing Tourism Research Globally

2010 ttra International Conference

# Nonresponse Bias in Internet-based Advertising Conversion Studies

Sangwon Park
*National Laboratory for Tourism & eCommerce School of Tourism & Hospitality Management Temple University*

Daniel R. Fesenmaier
*National Laboratory for Tourism & eCommerce School of Tourism & Hospitality Management Temple University*

**Nonresponse Bias in Internet-based Advertising Conversion Studies**

Sangwon Park
National Laboratory for Tourism & eCommerce
School of Tourism & Hospitality Management
Temple University

and

Daniel R. Fesenmaier
National Laboratory for Tourism & eCommerce
School of Tourism & Hospitality Management
Temple University

## ABSTRACT

*This study examines the extent of nonresponse bias in online advertising conversion studies. Two indicators (i.e., conversion rates and travel expenditure) assessing the tourism advertising effectiveness were compared using unweighted and weighted data sets. The results of this study using 24 locations throughout the U. S. confirm the conclusions of previous studies, showing consistent overestimates in advertising effectiveness. Several methodological and managerial implications of these findings are discussed.*

**Keywords:** *tourism advertising, conversion study, nonresponse bias.*

## INTRODUCTION

Tourism researchers have developed a variety of methods to assess the effectiveness of advertising campaigns (Kim, Hwang and Fesenmaier 2005; McWilliams and Crompton 1997; Siegel and Ziff-Levine 1990; Woodside and Ronkainen 1984). The accuracy and reliability of these approaches have been challenged in terms of sampling strategies, and nonresponse bias (Mok 1990; Woodside and Ronkainen 1984). Today, many tourism advertising studies use online surveys instead of traditional mailing surveys due to the benefits associated with the Internet including accessibility (Tierney 2000), low cost (Tse 1998), fast response (Weible and Wallace 1998), and verifiable delivery (James, Wotring, and Forrest 1995). With online surveys, one source for sampling bias has largely been eliminated as online-based surveys can be sent to the entire population rather than a sample of persons requesting travel information (Hwang and Fesenmaier 2004). However, Ellerbrock (1981) and Burke and Gitelson (1990) argued that nonresponse bias may overestimate the effectiveness of the tourism advertising (e.g, conversion rate and trip expenditure) because people who visited a destination are more likely to respond to a travel survey comparing to non-visitors. Some studies have suggested that the most effective approach to minimize nonresponse bias is by increasing response rate. However, it is argued that while increasing response rate is a good suggestion, it is not really practical due to the general patterns of low response rate in web surveys (Best et al. 2001; Kwak and Radler 2002). Also, despite high response rates, studies indicate that there still is the potential for significant differences between total sample and respondents (Bandilla, Bosnjak, and Altdorfer, 2003).

A number of studies published in the political, educational and medicine literatures suggest that various weighting procedures can be used to estimate the extent of nonresponse bias by adjusting the estimates provided by respondents to more closely represent the total target sample (Biemer and Lyberg, 2003). That is, because the true response of the population is unknown, adjusting the sample so that it "looks" like the population provides a viable way of "guessing" at the underlying behavior of the population. Based upon this approach, this study estimates the extent (mean and range) of nonresponse bias in online travel advertising conversion studies for twenty four destinations located throughout the United States. This paper followed a three step process in order to accomplish this goal. First, logistic regression was used to assess differences in respondents and nonrespondents in terms of their geographic and demographic characteristics. Next, the respondent data for each of the twenty four American destinations was *posthoc* weighted based on the geographic and demographic variables using two different weighting methods (post-stratification and propensity score weighting). Last, the estimates of conversion rates and travel expenditures for each of the twenty four destinations were compared between the unweighted data and two weighted data.

## LITERATURE REVIEW

For several decades, tourism researchers have sought to assess the effectiveness of tourism advertising by estimating the proportion of people responding to advertisements by actually visiting a destination and the amount of travel expenditure generated from these visits (Faulkner 1997). The conversion approach, a common approach used by tourism organizations, is most often based upon a direct response from those who requested information from the tourism organization. The advantages of the conversion approach include accessing to potential visitors, the straightforward implementation of estimation procedures, and the low cost of collecting data (Lankford, *et al.* 1995; McWilliams and Crompton 1997; Woodside and Sakai 2003). However, tourism researchers have identified several methodological deficiencies in the use of the conversion approach (e.g., Burke and Gitelson 1990; Hunt and Dalton 1983; Mok 1990; Siegel and Ziff- Levine 1994; Woodside 1981). The two main methodological problems that affect the validity and reliability of the advertising conversion estimates are: 1) sampling error (i.e., sampling precision and size), and 2) nonresponse error (Rylander, Propst, and McMurtry 1995; Woodside and Ronkainen 1984). However, the low cost of the Internet as a survey tool has largely eliminated the use of samples and therefore the problems associated with sampling error (Hwang and Fesenmaier, 2004).

Interestingly, nonresponse bias has become an even greater concern as response rates have declined substantially over the last decade and are often extremely low when using the Internet (Dolnicar, Laesser, and Matus 2009). Nonresponse bias occurs when " a significant number of people in the survey sample do not respond to the questionnaire and have different characteristics from those who do respond, when those characteristics are important to the study" (Dillman 2007, p 10). Burke and Gitelson (1990) argued that people who visit a destination will be more likely to respond to a survey than those people who did not visit. Therefore, nonresponse bias should lead to a significant ***inflation*** in the estimate of visitor conversion and expenditures. In other research, Woodside and Ronkainen (1984) argued that wave analysis (comparing the differences in demographic, attitudinal or behavioral variables across mail waves) is useful when information about nonrespondents is unavailable. Many studies have used wave analysis to identify nonresponse error including Lankford, Buxton, Hetzler and Little (1995), McCool (1991), and Woodside and Ronkainen (1984) and suggested this as a general approach to reduce the error. However, the use of this approach also has been challenged because it seems unreasonable to assume that late respondents are reliable substitutes for nonrespondents (Rylander, et al. 1995; Crompton and Cole 2001).

This study considers an alternative approach to asessing nonresponse bias in travel-related conversion studies. Following Makela (2003), it is argued that estimates of conversion rate and visitor expenditure can be estimated by the *posthoc* weighting of the response data. In this approach, nonresponse bias can be calculated by comparing conversion rates and visitor expenditure estimates using the "raw" response data and similar estimates made using weighted data. Post-stratification (i.e., a sort of population-based weighting) is, perhaps, the most widely used weighting method (Lessler and Kalsbeek 1992); this is where the weights are constructed to match the proportional distribution of strata concerning auxiliary variables (e.g., demographic and geographic) being available from both sample (i.e., response in this study) and population data (i.e., inquiries to the 24 DMOs). Post-stratification weighting enables researchers to "rebalance" the distribution of responses within the sample so as to correspond with the distribution (i.e., look like) of the overall population. This adjustment has been shown to significantly improve the accuracy of survey estimates by reducing bias (i.e., sampling and response errors) as well as increasing precision, especially for survey outcomes highly correlated with the post-stratifying variables (Little 1993).

Propensity score adjustments is another approach to *post hoc* weighting that has been shown to alleviate the confounding effects of the response mechanism (i.e., survey tools) by achieving a balance of covariates between the population and the sample (Rosenbaum and Rubin 1983a). Propensity scoring weight includes a strong theoretical foundation and has been extensively used within the statistical community (Lee and Valliant 2006; Rosenbaum and Rubin 1983b). A propensity score is the conditional probability of responding a survey and logistic regression is usually used to estimate the conditional probability based on the auxiliary information (refers to confounders) in this adjustment. The key strength of the propensity score weighting method is that it can be used to reduce a large set of confounding (or auxiliary) variables into a single propensity score (i.e., weight). That is, it addresses the challenge of integrating a series of independent variables such as age, gender, location and market investment into a single weighting scheme; recently, several studies have confirmed the usefulness of the propensity score weighting in online surveys (Lee 2006; Taylor 2000).

## METHODOLOGY

The data used in this study were obtained from two sources. First, the population data included <u>all</u> persons requesting information from the destination marketing organizations (DMOs) or the advertising firm of 24 destinations located throughout the United States during calendar year, 2009. The information known about every inquirer is limited to basic geographic and demographic attributes including: (1) Residence – measured using three different levels of closeness to the destination state; it is argued that based upon many studies, responses will differ significantly based upon location (Harzing 1997; Sheth and Roscoe 1975). A variable, IN-STATE was calculated to identify those inquirers that live in same state where the advertising is promoted; a second variable, ADJACENT STATE, was calculated to identify those inquirers living in bordering states; last OUT-STATES was calculated to identify those inquirers that live states further away; (2) TARGET MARKET – defined as whether or not (0/1) the inquirer resides within or outside of the markets defined by the advertising campaign; and, (3) PRIZM is a demographic segment –based upon the segmentation tool developed by Claritas, Inc., whereby each respondent is categorized into one of 66 demographic groups based on their five digit zipcode; it is argued that individuals that live near others are likely to have similar demographic and life styles.

The second data set is based upon an online survey of all persons included in the population data base. The online survey was distributed to 158,705 persons who requested brochure advertisements from the tourism organizations representing the 24 American destinations. All surveys were conducted during 2009 (i.e., from January to December in 2009) and applied the same web survey methodology including the same survey design, operation system, number of reminders, and amount of incentives. In particular, the survey used a three-step process: (1) the initial invitation was sent out along with the URL of the survey; (2) four days later, a reminder was delivered to those who had not completed the survey; and, (3) the final request for participation was sent out to those who had not completed the survey one week later. An Amazon.com gift card valued at $100 was provided to one winner for each campaign as an incentive for encouraging survey participation. The survey effort resulted in total 14,700 responses (an average response rate across the 24 campaigns of 9.58%).

The first step in the research process used logistic regression to evaluate the extent to which there are systematic differences between those that answered the conversion survey and those that did not. Logistic regression was deemed appropriate as the dependent variable is dichotomous indicating whether or not the respondent completed the survey (0 = No, 1 = Yes) and the independent variables are also indicators of geographic location and demographic makeup (i.e., residence state, target market and Prizm segmentation), and where it is argued that these variables are important behavioral factors effecting conversion rates and visitor expenditures (Couper et al., 1999; McWilliams and Crompton, 1997; Messmer and Johnson 1993; Sideman and Fik 2005). The specific form of the model is as follows:

$$= \beta_0 + In\ stateX_1 + Adjacent\ StateX_2 + PS1X_3 + \cdots + PS66X_{68} + InMarketX_{69}$$

where PS refers to the Prizm segmentation; Out-State in the residence state, PS 67 (‘unclassified’) in the Prizm segmentation and Outside-Market in the target market are used as reference groups.

Logistic regression was conducted for each of the 24 destinations, and the results show that respondents to the conversion studies for 22 of these campaigns are statistically significant; this finding indicates that there are significant behavioral differences between those that responded to the survey and those that did not. Please note that even though two studies (i.e., campaigns 3 and 10) showed no significant results, these campaigns were included in further analysis so that we could estimate the range of potential bias across all campaigns.

Two post hoc weighting approaches (i.e., post-stratification and propensity score weighting) using the respective variables (i.e., residence states, Prizm segmentation, and target market) were employed to adjust for the differences between respondents and nonrespondents. For the post-stratification, weights for each of the three sets of variables were calculated separately based upon the proportion of the total population (i.e., both respondents and nonrespondents) for each variable was divided by the proportion of the respondents for each relevant variable. A single weight, $W_i$, for each individual respondent was then calculated (Cordell, Betz, and Green 2002):

$$W_i = W_s * W_p * W_t$$

The same variables (i.e., target market, residence states, and Prizm segmentation) were also used to estimate weight values for propensity score adjustment following Rosenbaum & Rubin (1984). Specifically, logistic regression models were developed for each campaign and used to obtain the conditional probability (i.e., propensity score) of responding the online survey. Then, each respondent was classified into quintile based upon the propensity score. Studies indicate that creating five strata based on propensity scores remove approximately 90% of the removable bias (Cochran 1968; Rosenbaum & Rubin 1983).

## RESULTS

Conversion rates were estimated using the unweighted data and the two weighting schemes (post-stratification and propensity score weighting) and are presented in Table 1. The results show a substantial overestimation of the conversion rates when comparing the unweighted and weighted data. Specifically, the average conversion rate using the unweighted data is 41.3% while the average conversion rates for the weighted data are much lower at 34.8% (for the post stratification approach) and 37.0% (for the propensity weighting approach), respectively. There is substantial variation, however, in the accuracy of the weighting schemes. That is, there appear to be little differences in the estimates for at least 13 of the 24 campaigns; however, there are substantial differences in the estimates for at least 5 campaigns (i.e., Campaigns 5, 9, 14, 21 and 22). In Campaign 11, for example, the conversion rate using the propensity score weighting approach (91.3%) is only slightly higher than the estimates for the unweighted data (91.0%), and the post-stratification data (90.2%). However, the estimated conversion rate for Campaign 22 using the unweighted data is 57.1%; this estimate contrasts sharply with the 35.5% conversion rate based upon the weighted data using the post stratification approach (a difference of 21.6%) or the 44.76% using the propensity score weighting (a difference of 12.45). These differences in estimates can be explained by the failure to include additional factors (beyond target market, residence states and Prizm segmentation) that affect response behavior (i.e., completing the conversion survey).

The final set of analyses in this study focused on evaluating the "impact" of the response bias on conversion rates and visitor expenditure, arguing that they are often used as an important indicator of advertising effectiveness. Due to the survey structure of the expenditure question (i.e., trip expenditure was provided in 12 categories of differing ranges), the median value was used as an estimate of average trip expenditure (see Table 2). The results show that the unweighted median travel expenditure across all campaigns is $486, but this estimate is somewhat lower as compared to the estimates based upon the weighted data: $517 for the post stratification approach and $503 for the propensity score method. Comparison of campaigns shows that the median trip expenditure estimates for Campaigns 1, 12, 16, 21, 22, and 23 are underestimated, whereas the estimated median trip expenditure for those responding to Campaign 7 is overestimated by $20. The remaining 17 campaigns appear to have essentially the same median trip expenditure. However, when using the propensity score weighting approach, Campaigns 1, 7, 9, 12, and 23 are somewhat lower, while Campaigns 8 and 19 are over estimated by $100 and $130, respectively.

**Table 1**
**The Estimated Conversion Rates Between Unweight and Weighted Data Sets**

| Campaigns | Respondents | | | Unweighted | Post-Stratification | | Propensity Score Weighting | |
|---|---|---|---|---|---|---|---|---|
| | N | n | Response Rate (%) | Conversion Rate (%) | Conversion Rate (%) | Difference (%) | Conversion Rate (%) | Difference (%) |
| 1 | 8,453 | 791 | 9.4 | 21.2 | 19.2 | 2.0 | 20.2 | 1.0 |
| 2 | 9,908 | 933 | 9.4 | 52.2 | 48.9 | 3.3 | 50.4 | 1.8 |
| 3 | 1,931 | 131 | 6.8 | 32.8 | 32.4 | 0.4 | 32.2 | 0.6 |
| 4 | 5,193 | 410 | 7.9 | 33.2 | 29.8 | 3.4 | 29.3 | 3.9 |
| 5 | 5,374 | 326 | 6.1 | 37.7 | 25.2 | 12.5 | 29.8 | 7.9 |
| 6 | 11,403 | 819 | 7.2 | 40.2 | 37.5 | 2.7 | 37.3 | 2.9 |
| 7 | 6,054 | 328 | 5.4 | 19.8 | 12.2 | 7.6 | 15.4 | 4.4 |
| 8 | 7,574 | 769 | 10.2 | 43.0 | 33.7 | 9.3 | 36.9 | 6.1 |
| 9 | 7,849 | 565 | 7.2 | 60.2 | 44.3 | 15.9 | 50.7 | 9.5 |
| 10 | 1,721 | 291 | 16.9 | 70.1 | 68.2 | 1.9 | 68.9 | 1.2 |
| 11 | 1,858 | 189 | 10.2 | 91.0 | 90.2 | 0.8 | 91.3 | -0.3 |
| 12 | 5,513 | 420 | 7.6 | 53.6 | 50.6 | 3.0 | 50.2 | 3.4 |
| 13 | 4,744 | 360 | 7.6 | 36.4 | 34.8 | 1.6 | 32.7 | 3.7 |
| 14 | 5,209 | 441 | 8.5 | 42.2 | 31.4 | 10.8 | 34.3 | 7.9 |
| 15 | 4,031 | 392 | 9.7 | 41.6 | 37.0 | 4.6 | 39.9 | 1.7 |
| 16 | 9,888 | 773 | 7.8 | 53.7 | 45.0 | 8.7 | 48.3 | 5.4 |
| 17 | 5,339 | 389 | 7.3 | 19.5 | 17.5 | 2.0 | 19.9 | -0.4 |
| 18 | 3,257 | 319 | 9.8 | 33.5 | 28.2 | 5.3 | 29.3 | 4.2 |
| 19 | 8,458 | 825 | 9.8 | 17.8 | 14.3 | 3.5 | 15.1 | 2.7 |
| 20 | 9,044 | 652 | 7.2 | 25.9 | 22.4 | 3.5 | 24.2 | 1.7 |
| 21 | 7,843 | 696 | 8.9 | 43.8 | 22.7 | 21.1 | 30.5 | 13.3 |
| 22 | 14,766 | 1,036 | 7.0 | 57.1 | 35.5 | 21.6 | 44.7 | 12.4 |
| 23 | 3,744 | 333 | 8.9 | 37.5 | 29.7 | 7.8 | 31.7 | 5.8 |
| 24 | 9,551 | 820 | 8.6 | 26.8 | 24.8 | 2.0 | 24.5 | 2.3 |
| Minimum | 1,721 | 131 | 5.4 | 17.8 | 12.2 | 0.4 | 15.1 | -0.3 |
| Maximum | 14,766 | 1,036 | 16.9 | 91.0 | 90.2 | 21.6 | 91.3 | 13.3 |
| Mean | 6,613 | 542 | 8.5 | 41.3 | 34.8 | 6.5 | 37.0 | 4.3 |
| Standard Deviation | | | 2.2 | 17.5 | 17.6 | 6.1 | 17.2 | 3.7 |

As part of this analysis, total expenditure values were calculated for all inquirers of the campaign (i.e., number of inquirers x conversion rate x trip expenditure) with the idea of trying to understand the total impact of nonreponse bias. As can be seen in Table 2, the mean total revenue from visitors across the 24 campaigns is $136.3 million (based upon the unweighted data) and ranges from $22.2 million to $504.2 million. This estimate compares to a mean of $121.9 million using a post-stratification approach (a $14.4 million of mean difference) and $124.8 million using the propensity score weight approach (a $11.5 million of mean difference). Comparison of the campaigns shows that four campaigns (i.e., Campaign 1, 12, 16, and 23 using post-stratification) and six campaigns (i.e., 1, 9, 11, 12, 17, and 23 using propensity score weighting) are substantially "underestimated" (ranging from differences of $.14 million to $37.2 million). Also, the estimates for a number (7) of campaigns are substantially over estimated; for example, Campaign 22 is overestimated between $82.4 million (using propensity score weighting) and $91.1 million when using post stratification weighting.

## DISCUSSION

It is essential for destination marketing organizations to understand nonresponse bias when conducting the survey in order to accurately evaluate the effectiveness of the advertising campaign. This study identified significant nonresponse bias in Internet-based advertising study by comparing the estimated results of unweighted and weighted data based upon the results of 24 different advertising campaigns. Two types of weighting methods (i.e., post-stratification and propensity score weighting) were used to estimate the conversion rates and expenditure levels of the population of inquiries for each of the 24 campaigns in the United States. The results of the study indicate that the use of unweighted data to estimate advertising effectiveness leads to substantial and consistent over estimation of conversion rate, but there is limited "bias" in the estimates of median visitor expenditures. These estimates appear to lead to a substantial over estimation of the overall value of the campaign. Indeed, the results of the 24 American studies indicate that conversion studies using unweighted data leads to an overall over estimate of approximately 10%.

Of course, there is substantial variation in sign (over or under estimates) and extent of bias. Indeed, as expected based upon the distributional assumptions underlying this approach, the effect of response bias seems to be quite limited for many of the campaigns. In this study the conversion rate estimates of approximately half of the campaigns were within ± 3% of that based upon the unweighted data; concomitantly, there were 5 – 6 campaigns with very large errors (ranging from 10% - 22%). In this latter case, it seems that this is the results of the failure to include appropriate variables in the weighting scheme; therefore, it appears that additional variables should be considered when developing alternative weighting schemes in order to improve substantially the quality of the estimates. These variables might include other behavioral variables such as knowledge and image of the destination, competition-related variables (i.e., the number and competitiveness of alternative nearby destinations) and various aspects (i.e., target markets, amount of investment, etc.) of the specific campaign.

Although there are several limitations in this study that may influence the results, this study of 24 different American tourism campaigns provides a significant understanding in the nature (mean and range) of impact of nonresponse bias in conversion studies. Additionally, in the case where it is hard to obtain a reference survey in the advertising study, it appears that the two weighting methods used in this study can be useful in assessing the errors in the response data. Last, it is hoped that this study provides the basis for additional studies which incorporate information relative to inquirers and the advertising programs and further consideration of ways to better manage possible bias in conversion surveys.

**Table 2**
**The Estimated Travel Expenditure Between Unweight and Weighted Data Sets**

| Campaigns | | Unweighted | | | Post-Stratification | | | | Propensity Score Weighting | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Conversion Rate (%) | Median ($) | Total Expenditure (million $) | Conversion Rate (%) | Median ($) | Total Expenditure (million $) | Difference (million $) | Conversion Rate (%) | Median ($) | Total Expenditure (million $) | Difference (million $) |
| 1 | 8,453 | 21.2 | 650 | 116.5 | 19.2 | 900 | 146.1 | -29.6 | 20.2 | 900 | 153.7 | -37.2 |
| 2 | 9,908 | 52.2 | 900 | 465.5 | 48.9 | 900 | 436.1 | 29.4 | 50.4 | 900 | 449.4 | 16.1 |
| 3 | 1,931 | 32.8 | 350 | 22.2 | 32.4 | 350 | 21.9 | .27 | 32.2 | 350 | 21.8 | .41 |
| 4 | 5,193 | 33.2 | 900 | 155.2 | 29.8 | 900 | 139.3 | 15.9 | 29.3 | 900 | 137.0 | 18.2 |
| 5 | 5,374 | 37.7 | 250 | 50.7 | 25.2 | 250 | 33.9 | 16.8 | 29.8 | 250 | 40.0 | 10.6 |
| 6 | 11,403 | 40.2 | 1,100 | 504.2 | 37.5 | 1,100 | 470.4 | 33.9 | 37.3 | 1,100 | 467.9 | 36.4 |
| 7 | 6,054 | 19.8 | 370 | 44.4 | 12.2 | 350 | 25.9 | 18.5 | 15.4 | 450 | 42.0 | 2.4 |
| 8 | 7,574 | 43.0 | 350 | 114.0 | 33.7 | 350 | 89.3 | 24.7 | 36.9 | 250 | 69.9 | 44.1 |
| 9 | 7,849 | 60.2 | 250 | 118.1 | 44.3 | 250 | 86.9 | 31.2 | 50.7 | 350 | 139.3 | -21.2 |
| 10 | 1,721 | 70.1 | 250 | 30.2 | 68.2 | 250 | 29.3 | .82 | 68.9 | 250 | 29.6 | .52 |
| 11 | 1,858 | 91.0 | 250 | 42.3 | 90.2 | 250 | 42.0 | .37 | 91.3 | 250 | 42.4 | -.14 |
| 12 | 5,513 | 53.6 | 250 | 73.9 | 50.6 | 350 | 97.6 | -23.8 | 50.2 | 350 | 96.9 | -23.0 |
| 13 | 4,744 | 36.4 | 550 | 95.0 | 34.8 | 550 | 90.8 | 4.2 | 32.7 | 550 | 85.3 | 9.7 |
| 14 | 5,209 | 42.2 | 350 | 76.9 | 31.4 | 350 | 57.3 | 19.7 | 34.3 | 350 | 62.5 | 14.4 |
| 15 | 4,031 | 41.6 | 350 | 58.7 | 37.0 | 350 | 52.2 | 6.5 | 39.9 | 350 | 56.3 | 2.4 |
| 16 | 9,888 | 53.7 | 350 | 185.8 | 45.0 | 450 | 200.2 | -14.4 | 48.3 | 350 | 167.2 | 18.7 |
| 17 | 5,339 | 19.5 | 700 | 72.9 | 17.5 | 700 | 65.4 | 7.5 | 19.9 | 700 | 74.4 | -1.5 |
| 18 | 3,257 | 33.5 | 250 | 27.3 | 28.2 | 250 | 23.0 | 4.3 | 29.3 | 250 | 23.9 | 3.4 |
| 19 | 8,458 | 17.8 | 700 | 105.4 | 14.3 | 700 | 84.7 | 20.7 | 15.1 | 570 | 72.8 | 32.6 |
| 20 | 9,044 | 25.9 | 700 | 164.0 | 22.4 | 700 | 141.8 | 22.2 | 24.2 | 700 | 153.2 | 10.8 |
| 21 | 7,843 | 43.8 | 450 | 154.6 | 22.7 | 550 | 97.9 | 56.7 | 30.5 | 450 | 107.7 | 46.9 |
| 22 | 14,766 | 57.1 | 450 | 379.4 | 35.5 | 550 | 288.3 | 91.1 | 44.7 | 450 | 297.0 | 82.4 |
| 23 | 3,744 | 37.5 | 250 | 35.1 | 29.7 | 350 | 38.9 | -3.8 | 31.7 | 350 | 41.5 | -6.4 |
| 24 | 9,551 | 26.8 | 700 | 179.2 | 24.8 | 700 | 165.8 | 13.4 | 24.5 | 700 | 163.8 | 15.4 |
| Minimum | 1,721 | 17.8 | 250 | 22.2 | 12.2 | 250 | 21.9 | -29.6 | 15.1 | 250 | 21.8 | -37.2 |
| Maximum | 14,766 | 91.0 | 1100 | 504.2 | 90.2 | 1100 | 470.4 | 91.1 | 91.3 | 1100 | 467.9 | 82.4 |
| Mean | 6,613 | 41.3 | 486 | 136.3 | 34.8 | 517 | 121.9 | 14.4 | 37.0 | 503 | 124.8 | 11.5 |
| SD | | 17.5 | 249.6 | 132.0 | 17.6 | 253.1 | 120.5 | 24.9 | 17.2 | 250.6 | 120.7 | 24.9 |

Note: Total expenditure was calculated by (number of inquirers) x (conversion rates) x (Median); Difference was calculated by (the total expenditure from unweight) – (the total expenditure from each weighting method)

## REFERENCES


Bandilla, W., Bosnjak, M., and Altdorfer, P. (2003). "Survey Administration Effects? A Comparison of Web-Based and Traditional Written Self-Administered Surveys Using the ISSP Environmental Module." *Social Science Computer Review*, 21; 235-243.

Best, S. J., B. Krueger, C. Hubbard, and A. Smith (2001). "An Assessment of the Generalizability of Internet Surveys." *Social Science Computer Review*, 19 (2): 131-45.

Biemer, P.P. and L.E. Lyberg (2003). Introduction to Survey Quality. New York: John Wiley.

Burke, J. F., and R. Gitelson. (1990). "Conversion Studies: Assumptions, Applications, Accuracy And Abuse." Journal of Travel Research, 28: 46-50.

Cordell, H. K., C. J. Betz, and G. T. Green (2002). "Recreation and the Environment as Cultural Dimensions in Contemporary American Society" *Leisure Sciences*, 24 (1): 13-42.

Couper,M. P., J. Blair, and T. Triplett (1999). "A Comparison of mail and e-mail for a Survey of Employees in U.S. Statistical Agencies." *Journal of Official Statistics*, 15: 39-56.

Crompton, J. L. and S-T. Cole (2001). "An Analysis of 13 Tourism Surveys: Are Three Waves of Data Collection Necessary?" *Journal of Travel Research*, 39(4):356-368

Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method.* 2nd ed. New York: John Wiley.

Dolnicar, S., C. Laesser, and K. Matus (2009). "Online Versus Paper: Format Effects in Tourism Surveys" *Journal of Travel Research*, 47(24): 295-316.

Ellerbrock, D. (1981), "Improving Coupon Conversion Studies," *Journal of Travel Research*, 19(4), 37-38.

Faulkner, B. (1997). "A Model for the Evaluation of National Tourism Destination Marketing Programs." *Journal of Travel Research*, 35: 23 - 32.

Harzing, A-W. (1997). "Response rates in International Mail Surveys: Results of a 22-Country Study." *International Business Review*, 6 (6):641-65.

Hwang, Y. H., and D. R. Fesenmaier (2004). "Coverage Error Embedded In Self-Selected Internet-Based Samples: A Case Study of Northern Indiana." *Journal of Travel Research*, 42(3): 297-304.

Hunt, J. D. and M. J. Dalton (1983). "Comparing Mail and Telephone for Conducting Coupon Conversion Studies." *Journal of Travel Research*, 21: 16 - 18.

James, M. L., C. E. Wotring, and E. J. Forrest (1995). "An Exploratory Study of the Perceived Benefits of Electronic Bulletin Board Use and Their Impact on Other Communication Activities." *Journal of Broadcasting and Electronic Media*, 39 (1): 30-50.

Kim, D-Y., Y-H. Hwang, and D. R. Fesenmaier. (2005). "Modeling Tourism Advertising Effectiveness." *Journal of Travel Research*, 44: 42-9.

Kwak, N., and B. Radler (2002). "A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality." *Journal of Official Statistics*, 18(2): 257-273.

Lankford, S. V., B. P. Buxton, R. Hetzler, and J. R. Little (1995). "Response Bias and Wave Analysis of Mailed Questionnaires in Tourism Impact Assessments." *Journal of Travel Research*, 33: 8-13.

Lee, S. (2006). "Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys" *Journal of official statistics*, 22(2): 329–349

Lee, S., and Valliant, R. (2008). "Post-Survey Weighting Methods Using Propensity Scores: A Review," chapter 8 in Advances in Telephone Survey Methodology, John Wiley.

Lessler, J. T., and W D. Kalsbeek (1992). Nonsampling Error m Surveys. New York: John Wiley and Sons.

Little, R. J. A. (1993). “Post-Stratification: A Modeler’s Perspective” *Journal of the American Statistical Association,* 88(423): 1001-1012.

Makela, P. (2003). “Impact of Correcting for Nonresponse by Weighting on Estimates of Alcohol Consumption.” *Journal of studies on alcohol*. 64(4): 589-96.

McCool, S. F. (1991). “Using Probabilistic Incentives to Increase Response Rates to Mail-return Highway Intercept Diaries.” *Journal of Travel Research,* 30(2): 17-19.

McWilliams, E. G., and J. L. Crompton (1997). “An Expanded Framework for Measuring the Effectiveness of Destination Advertising.” *Tourism Management*, 18(3): 127-37.

Messnmer, D. J., and R. R. Johnson (1993). “Inquiry Conversion and Travel Advertising Effectiveness.” *Journal of Travel Research,* 31(4): 14-21.

Mok, H. M. K. (1990). “A Quasi-Experimental Measure Of The Effectiveness Of Destinational Advertising: Some Evidence From Hawaii.” *Journal of Travel Research,* 29: 30-4

Rosenbaum, P. R., and D. B. Rubin (1983). **“**The central role of the propensity score in observational studies for causal effects” Biometrika, 70(1): 41-55.

Rosenbaum, P. R., and D. B. Rubin (1983). **“**Reducing Bias in Observational Studies Using Subclassification on the Propensity Score” Journal of American Statistical Association, 79(387): 516-524.

Rubin, D. B. (1997). “Estimating Casual Effects from Large Data Sets Using Propensity Scores.” *American College of Physicians,* 127(8) :757-763.

Rylander, R. G., Propst, D. B., and Mcmurtry, T. R. (1995). “Nonresponse and Recall Biases in a Survey of Traveler Spending.” *Journal of Travel Research,* 33(4): 39-45.

Sheth, J. N., and A, M. Roscoe (1975). “Impact of Questionnaire Length, Follow-up Methods, and Geographical Location on Response Rate to a Mail Survey.” *Journal of Applied Psychology,* 60 (2): 252-254.

Sideman, C. F. and T. J. Fik (2005). “Modeling Spatial Patterns of Recreational Boaters: Vessel, Behavioral, and Geographic Considerations” *Leisure Sciences*, 27: 175–189.

Siegel, W., and W. Ziff-Levine (1990). “Evaluating Tourism Advertising Campaigns: Conversion vs. Advertising Tracking Studies.” *Journal of Travel Research,* 28: 51-5.

Taylor, H. (2000). “Does Internet Research “work”? Comparing on-line Survey Results with Telephone Surveys.” *International Journal of Market Research*, 42(1), 51-63.

Tierney, P. (2000). “Internet-Based Evaluation of Tourism Web Site Effectiveness: Methodological Issues and Survey Results.” *Journal of Travel Research,* 39: 212-19.

Tse, A. C. B. (1998). “Comparing the Response Rate, Response Speed and Response Quality of Two Methods of Sending Questionnaire: E-Mail vs. Mail.” *Journal of Marketing Research Society*, 40 (4): 353-61.

Weible, R., and J. Wallace (1998). “Cyber Research: The Impact of the Internet on Data Collection.” *Marketing Research*, 10 (3): 19-26.

Woodside, A. G. (1981). "Measuring the Conversion of Advertising Coupon Inquiries into Visitors." *Journal of Travel Research,* 19: 38-41.

Woodside, A. G., and I. A. Ronkainen (1984). “How Serious Is Nonresponse Bias in Advertising Conversion Research?” *Journal of Travel Research,* 22: 34-37.

Woodside, A. G., and M .Y. Sakai (2003). “Meta-Evaluation: Assessing Alternative Methods of Performance Evaluation and Audits of Planned and Implemented Marketing Strategies.” *Advances in Business Marketing and Purchasing,* 12: 549-663.