
Jay Beaman

Colorado State University
Auctor Consulting Associates, Ltd.

Jerry J. Vaske

Human Dimensions in Natural Resources Unit
Department of Natural Resource Recreation and Tourism
Colorado State University

Reliability, Accuracy and Bias in Survey Estimates in Tourism Surveys

Reliability, accuracy, and bias are common concepts that are sometimes misunderstood and used inappropriately. This note (1) clarifies the distinctions between reliability, accuracy and bias for *Estimates (ê) Based on Observable Behaviors* (EBOB), and (2) illustrates the importance of recognizing bias in them. Estimates considered include proportions in categories and totals or means of numbers/ amounts. Examples of estimate bias for large professionally designed and run surveys are presented. It is argued that, rather than treating reliability as accuracy, an estimate being *adequately* close to its correct value *should* be an important consideration in using it in tourism planning and management decision making.

Jay Beaman

Adjunct Faculty, Colorado State University
Auctor Consulting Associates, Ltd.
Cheyenne, WY 82009

Jerry J. Vaske

Human Dimensions in Natural Resources Unit
Department of Natural Resource Recreation and Tourism
Colorado State University
Fort Collins, CO 80523

Dr. Beaman has academic qualifications in physics, mathematics, statistics and sociology (methodology, demography and planning). His has worked as a researcher for over 4 decades and was the director of socio-economic research for Parks Canada for 20 years. He has published numerous articles. He serves on editorial boards of some journals and is a reviewer for several tourism and other journals.

Jerry J. Vaske is a professor in the Department of Natural Resource Recreation and Tourism at Colorado State University, and a co-leader of the Human Dimensions in Natural Resources Unit. He received his B.A. and M.A. from the University of Wisconsin, and his Ph.D. from the University of Maryland. Jerry's areas of research interest include survey methodology and analysis. He is co-editor of the journal *Human Dimensions of Wildlife*.

Introduction

Reliability, accuracy, and bias are common concepts in tourism research. Unfortunately, these concepts are sometimes misunderstood and used inappropriately. This research note has two objectives: (1) to clarify the distinctions between reliability, accuracy and bias for *Estimates* (\hat{e}) *Based on Observable Behaviors* (EBOB), and (2) to illustrate the importance of recognizing bias in survey research EBOB. EBOB include proportions computed based on being in categories (proportion female or VFR traveler) and totals or means of numbers/ amounts (e.g., number of times doing “x”, money/ time spent on “y” doing/during “z”). Based on behavior being observable, EBOB are ratio scale variables with a meaningful zero point (nobody in category “x”, no participation, no expenditure), and they have a correct value (e) for an estimate (\hat{e}). In practical terms, an estimate being *adequately* close to its correct value (e.g., \hat{e} within 10% of e), *should* be an important consideration in using it in tourism planning and management decision making.

Definitions

Bias

In mathematical statistics an estimate (\hat{e}) is biased if its expected value ($E[\hat{e}]$) does not equal its correct value (e).

Definition 1: An EBOB is biased if $\beta = E[\hat{e}] - e \neq 0$.

Problems arise in interpreting what “being biased” means because *bias* is used in different senses. First, bias can refer to a respondent intentionally giving misleading answers. Social desirability, acquiescence, cognitive demand effects, response order effects, and survey mode (e.g., mail versus telephone) have all been suggested as potential causes of intentional bias. Second, responses being biased can refer to some respondents not recalling the correct answer to

a question (Krosnick, 1999). Overuse of 0-5 ending approximate responses results in 0-5 response heaping (Vaske, Huan, & Beaman, 2003). Bias in estimates made using data with 0-5 heaps can be a consequence of how researchers calculate estimates (Beaman, Vaske, & Grenier, 1998). Using the right estimator matters (Sheaffer, Beaman, O'Leary, Williams, & Mason, 2000). Moreover, even if all respondents give correct responses, non-response bias occurs when estimates for the population would differ from estimates for respondents (Beaman, Vaske, & Miller, in press).

Reliability

Reliability is associated with reproducibility. The standard deviation of an EBOB, $\sigma(\hat{\epsilon})$, is a measure of its reproducibility. An operational definition of reliability is therefore:

Definition 2: An $\hat{\epsilon}$ has a given level of reliability if the absolute value of the ratio

$$|100(\hat{\epsilon} - E[\hat{\epsilon}]) / E[\hat{\epsilon}]| \text{ is less than some appropriate percent.}$$

The *appropriate* percent is use dependent. For some uses, an estimate might be considered acceptably reproducible if it is *usually* within 10% (i.e., $|100(\hat{\epsilon} - E[\hat{\epsilon}]) / E[\hat{\epsilon}]| < 10\%$ with 95% probability). In another situation, being reliable could imply within 1% with 99% probability. For a small sample size, $\hat{\epsilon}$ can frequently differ from $E[\hat{\epsilon}]$ by 50% while for very large samples it can rarely vary from $E[\hat{\epsilon}]$ by more than 1% (Beaman, Huan & Beaman, 2004). When estimate variability is given as the coefficient of variation (e.g., Canadian Travel Survey, CTS) one only has reliability information *unless bias is zero* (Statistics Canada, 1997). Unfortunately, knowing the reliability of an estimate says nothing about its accuracy *when bias is unknown*.

Accuracy

Intuitively, *accurate* refers to being close to the correct value.

Definition 3: \hat{e} is acceptably accurate if there is a given probability (e.g., 95% certainty) that $|(\hat{e} - e)/e|$ is less than some percent Δ .

What Δ is appropriate (e.g., 5%) depends on how an estimate will be used. Researchers regularly claim that an estimate is accurate within $\pm 5\%$ based on some value of $\sigma(\hat{e})$. If an estimate is biased, this is not correct since $\sigma(\hat{e})$ is measured about $E(\hat{e})$. When $|(E(\hat{e}) - e)/e| = |\beta|/e$ is too large to have, e.g., a 95% probability of estimates being within 5% of e , then \hat{e} cannot be accurate “within 5%”.

Tourism Survey Estimates and Bias

Tourism survey estimates being seriously biased is rarely considered. Still, statements about accuracy often rest on $\sigma(\hat{e})$ decreasing with sample size; and the assumptions that “responses are unbiased” and that “non-respondents are like respondents.” Kunert (1998), however, has demonstrated that professionally run German travel surveys with about 40,000 respondents have serious bias problems. Changes in the survey design and procedures resulted in seriously biased estimates. Research has also show that Canadian Travel Survey trend estimates for 1996 and 1997 which have a low probability of $\sigma(\hat{e})/E(\hat{e})$ being over 2% were in error by far more than this because of bias (Beaman, et al., 2001). For Illinois waterfowl hunting, estimates based on a widely-used “traditional” survey method were found to be biased by about 15% (Beaman et al., in press). Biases in fishing estimates for Colorado have been estimated at about 7%, unless a correction for 0-5 heaping is made (Beaman et al., 1997). Given that for most survey estimates there is no *direct measurement of or correction for* bias from under- or over-reporting or from non-response, claims of accuracy are regularly based on assumptions.

Some researchers have modified survey methodology to address bias concerns. For the National Fishing, Hunting and Wildlife-Associated Recreation (FHWAR) survey, the length of

the respondents' recall period was reduced (Chu et al., 1992). Unfortunately, reducing bias can cause problems in interpreting estimates. For example, Miller and Anderson (2001) sent half of the hunters in a sample a "hunting activity record-card" for logging their hunting activities (e.g., days spent hunting, harvest on those days). Because hunters who received the record-card were less likely to provide 0-5 ending responses, using record-cards was concluded to yield more accurate responses. Survey methodology was changed resulting in a 15-20% discontinuity in estimates. Further research showed estimates primarily changed as a result of more reporting of low days of participation and harvest by record-card recipients (Beaman et al., in press). Phrasing in survey change and adequate examination of causes of bias are important to being able to compare estimates.

Objective assessment of the accuracy of EBOB involves comparisons of estimates to standards (Beaman, 2001). Chase and associates (Chase & Harada, 1984), for example, linked participation records of members maintained by a health club with self-reported participation. Self-reported participation was substantially higher than participation based on use records. For a survey like FHWAR for which it is possible to estimate the numbers of certain types of licenses sold, a standard is the actual number sold. For Wyoming for resident fishing, comparison resulted in an estimate that was 20% low with a probability < 1% that the difference occurred by chance.

Conclusion

Although large sums of money are spent on surveys such as the FHWAR and CTS, some estimates based on these surveys are biased by 15-20%. Although some estimates may be *appropriately* accurate for the uses made of them, diverting some of the dollars spent on data collection to determine bias and to developing bias correction factors seems prudent. Given that

bias can be large, without knowledge of the magnitude and direction of bias, treating reliable estimates as accurate is not wise. Hopefully, tourism research has advanced to the point where researchers will pay serious attention to obtaining estimates that meet not only reliability criteria, but accuracy criteria as well.

References

- Beaman, J. G. (2001). Comment on Digit preference in reported harvest among Illinois waterfowl hunters by Craig A. Miller and William L. Anderson. *Human Dimensions of Wildlife*, 7(1), 67-72.
- Beaman, J. G., Vaske, J. J., & Grenier, M. (1998). A prototype model for estimating and correcting bias in digit preference / number preference. *Tourism Analysis*, 2(2), 77-90.
- Beaman, J. G., Vaske, J. J., & Miller, C. (in press). Improving the accuracy of survey estimates. *Human Dimensions of Wildlife*.
- Beaman, J. G., Beaman, J. P., O'Leary, J. T., & Smith, S. (2001). The impact of seemingly minor methodological changes on estimates of travel and correcting bias. *Tourism Analysis*, 5, 91-96.
- Chase, D. R., & Harada, M. (1984). Response error in self-reported recreation participation. *Journal of Leisure Research*, 16(4), 322-329.
- Chu, A., Eisenhower, D., Hay, M., Morganstein, D., Neter, J., & Waksberg, J. (1992). Measuring the recall error in self-reported fishing and hunting activities. *Journal of Official Statistics*, 8(1), 19-39.
- Krosnick, J. A. (1999). Survey Research. *Annual Review of Psychology*, 50, 537-567.

- Kunert, U. (1998). Detecting Long-term Trends in Travel Behaviour: Problems Associated with Repeated National Personal Travel Surveys (pp. 263-277). In *Travel Behaviour Research: Updating the State of Play*, J. Ortúzar, D. Hensher, & S. Jara-Díaz (eds.). Oxford: Elsevier Science Ltd.
- Sheaffer, A. L., Beaman, J., O'Leary, J. T., Williams, R. L., & Mason, D. M. (2000). Weighting issues in recreation research in identifying support for resource conservation management alternatives. In G. Kyle (Ed.). *Proceedings of the 2000 Northeastern Recreation Research Symposium* (pp. 183-186). Bolton Landing, NY.
- Miller, C. A., & Anderson, W. L. (2001). Digit Preference in Reported Harvest among Illinois Waterfowl Hunters. *Human Dimensions of Wildlife*, 7 (1): 55-65.
- Statistics Canada (1997). *CTS MicroData File Documentation Manual*. Ottawa: Statistics Canada.
- Vaske, J. J. Huan, T. C., & Beaman, J. (2003). The use of multiples in anglers' recall of participation and harvest estimates. *Leisure Sciences*, 25, 399-409.