Surveys: What Does One Need to Know More About Them?

By
Harjanto Djunaidi[1]

A Selected Paper

Southern Agricultural Economics Association Annual Meetings
Orlando, Florida
February 4-8, 2006

[1]Assistant Professor of Agribusiness, School of Agribusiness and Agriscience. Middle Tennessee State University at Murfreesboro, TN. Please direct comments on this article to djunaidi at djunaidi@mtsu.edu.

Introduction

Agricultural economists are more concerns on how to perform statistical analyses to answer research questions. These preferences can be found in Agricultural Economics graduate programs through out the country. On Average, three to four classes in statistical analysis, mathematical statistics or econometrics classes are required before a student can earn a PhD degree in Agricultural Economics. What missing in the Agricultural Economics curricula is lack of discussions on the process of data gathering through surveys.

The lack of discussions on data gathering is understandable since most Agricultural Economists in the past have dealt mostly with historical and publicly available data published by the government institutions under the umbrella of United States Department of Agriculture or from other public and non-profit organization such as the FAO. The use of surveys as data collection method by Agricultural Economisis is not new and has been used in the past. However, in more recent years, there were clear indications that its popularity has increased as a means to collect information, especially in consumers' attitude or willingness to pay studies and others. This trend can be found in recent publications in the areas of Agricultural Economics as well as in Agribusiness. The second factor that has caused surveys to be more popular than they were in the past as a major tool to collect data is the availability of computer software used in other areas such as in marketing research to assess consumers' purchasing behavior. For example, the use of conjoint analysis is becoming a popular method in the profession to study opinion, attitude or preference toward a policy, government programs, or alternative

choices in studying market competition, product development or agrimarketing products in general.

Past curricula did not prepare Agricultural Economics students to design and write a set of good surveys questions. Students also never been taught on how to measure and recognize good surveys from the bad ones. As a result, most surveys conducted in the profession perhaps are questionable and inadequate. Most researchers in the Agricultural Economics profession use the approach for data collection without recognizing certain or preliminary steps that need to be fulfilled before conducting the surveys. This situation can be explored from what have been published in the Agricultural Economics or Agribusiness journals. Reliability of the surveys was rarely reported in most, if not all studies that have been done in the past. This fact suggests that something is missing and needs to be corrected so that a more appropriate way of doing the surveys can be done in the future.

This paper discusses several important steps that need to be pursued to make sure that overall questions asked in the surveys fulfill internal consistency reliability (Cronbach, 1951). Inconsistent measure of behavior, preference or attitude due to a poorly design surveys will increase measurements errors (ME) as measured by increasing standard deviation or variance (Crocker and Algina, 1986). From the econometrics or estimation point of view ME affects the use of biased estimator by researchers in their study. If this happens, it will cause the estimated coefficient to be far from its true value and will affect the results of any made inference (Green). If this is the case, it means the researcher just used a proxy variable instead of a latent variable in her or his analyses or econometrics estimation (Maddala). Another important problem that can occur from

estimating statistical models with error in the variable is related with consistency issues, especially in a large sample (Mood, Graybill and Boes). ME, efficiency as well as consistency issues are important in most econometrics estimation process. However, the problem becomes more crucial in the Agricultural Economics profession because of the policy implications or recommendations made following the completion of the study. Therefore, agricultural economists need to be aware that such errors can not be ignored and appropriate steps in surveys design need to be taken into account properly.

Objectives:

The objective of this study is to examine what does a researcher need to know before conducting surveys to collect data as one of the initial step in her or his research project. It is particularly important in the Agricultural Economics profession to address this problem since both past and mostly current curricula in the Agricultural Economics graduate program do not address or discuss such important issues. Instead, the profession is more concern on methods to analyze the data once they are collected. Unfortunately, current agricultural economics courses do not teach the students on how to write appropriate questions for data collection purposes from a group of respondents or population. As a result, the conclusions and policy recommendations of such a study could be over or under stated the real situations.

Data and Methods

Surveys questions are asked to measure certain opinion, preference or attitude of a group of respondents on certain issues. So, the surveys need to be designed such that the questions measure what it supposed to measure. To know that this objective is met, one

can calculate the internal consistency reliability as stated by Cronbach. His seminal work helps designing and collecting data to study the attitude or opinion of a group of people who got affected by policies or regulations or a group of people toward a product. This paper addresses these issues and shows how poorly design surveys could result in misleading conclusions.

Other professions such as marketing researchers or psychologists have used the reliability test extensively in scale development or attitude assessment study in order to minimize ME. Internal consistency which is also known by alpha coefficient measures the overall reliability of items asked in the surveys. The coefficient of alpha is estimated from the following equation:

Equation(1). $$\alpha = \frac{k}{(k-1)}\left[1 - \frac{\sum \sigma_i^2}{\sigma_X^2}\right]$$

Where:

k = The number of items.
$\sum \sigma_i^2$ = The sum of variance on item i.
σ_X^2 = Total variance of items asked.

Items consist of a set of questions used to assess respondent's attitudes, behaviors, or preferences on particular issues. The magnitude of alpha is between zero and one. The closer alpha to one could be interpreted as the more reliable the study is because the survey questions measure what they supposed or intended to measure.

Dillman (2000) mentioned four possible errors that one might potentially do in conducting surveys. These four common errors are: (1). Sampling error; (2). Coverage error; (3). Measurement error; and (4). Non responses error. Among these four types of

errors, measurement errors are perhaps the most crucial one for reasons that have been discussed above. Measurement errors (ME) occur because of poor questions wording or presented in such away that inaccurate or un-interpretable answer are obtained (Dillman). The other three errors have been discussed or covered well in most econometric texts. Therefore, they will not be discussed in detail. Crocker and Algina argued that ME are the deviations from the true scores or the "true traits" or "the true constructs" or the "true responses" given by the respondents. This deviation in econometric is known as standard deviation or the square of it as measured by the variance. The deviation in the context of ME could be positive or negative. Therefore, the larger the value of ME can be interpreted as the further away the researcher of getting the true traits or constructs or answers from the surveyed respondents. In other words, if the ME is significantly large, then the study is useless because the surveyed questions measure something else than what they are intended to measure which is totally inaccurate. The consequences of ME are just the same with the estimated variance in the econometric estimation procedures. A large variance can affect both the unbiasedness and efficiency of the estimator in a small sample case. Therefore, the estimator might not fulfill the following two conditions for UMVUE (uniformly minimum-variance unbiased estimator):

(1). $E_\theta[T^*] = \tau(\theta)$ for all θ in Φ; and
(2). $Var_\theta[T^*] \leq Var_\theta[T]$
Where:
T^* is an estimator or sample statistics randomly drawn from a population;
$\tau(\theta)$ = an estimate (true parameter) which value depends on parameter θ.
Var is the corresponding variance.
Φ is the parameter space.

Given the above explanations, the measurement errors have serious consequences toward getting a representative estimator or the sample statistics which satisfies the

UMVUE desirable properties. Crocker and Algina further argued that there are two sources of ME. The first is called systematic errors and the second one is the random errors. The systematic errors are errors which consistently affect an individual's observed score because of the characteristics of the individual or test that is not related to the construct being measured. An example of this type of errors is that respondents always answer "disagree" if a question seems to be ambiguous or vague. On the other hand, random errors are errors affecting an individual score or choice of options which happens purely by chance. For example, a respondent may not feel well or is sick when answering a questionnaire.

Both errors are important to be addressed and could affect the results of getting the representative estimator. While a researcher should minimize the random errors, they definitely need to take significant efforts to avoid the systematic errors because this type of errors is under her or his control. As previously mentioned one of the sources where systematic errors could occur is coming from poor survey design or asked questions. Surveys questions can be seen as efforts of trying to draw sample from the population such that useful statistics (estimator) can be generated. Therefore, poor surveys will not generate a representative estimator (either sample means or variance) which is close to the true value of the population means or variance.

Dillman classified two types of researchers when dealing with writing a survey questions. The first researcher is the one who not even know what she or he wants to know from the respondents, except in a general sense. On the other hand, the second type has something to do with a situation where the researcher knows what he or she wants but having a difficult time to come out with the "right questions" to ask. Dillman

suggested asking critical questions which can guide a researcher to accomplish her or his research project. This paper will use one example and show readers on how important a right question to be asked on surveys. An experiment was conducted by a group of students on bread buying purchase decision. An experiment was conducted by a group of students to study bread purchase decision. This activity is part of the class project that students have to accomplish. Set of surveys questions were constructed in class with students' participation. They were asked to identify important attributes that affect bread buyers in their purchase decision. The population of this study is the agriculture students at MTSU. The following question was asked by the surveyors to their peers in the experiment:

Experiment question #1: Which of the following factors do you consider when you buy the bread?

- ☐ Price
- ☐ Softness / Freshness
- ☐ Nutrition content
- ☐ Packaging
- ☐ Convenience / Availability

This question does not guarantee that respondents will choose one attribute out of the five choices given in the surveys due to the way the question is phrased. The use of the word "factors" in the plural form makes the respondents think that they could or have to choose more than one choice of attributes. When the results of the experiment were tallied, the students realized that many of the respondents chose different combination of those five factors. These combinations certainly will affect the probability of a certain choice being picked by the respondents. This type of systematic errors surely will increase the ME which in turns will affect the expected value, variance as well as the

results of the study. The question could be fine if the respondents were asked to choose only one choice which reflects their best choice.

In the example above a set of choices that the respondents have been asked is not ordered. The situation will be more complex if the order is important such as in a study where a Likert scale is used. Either the order is or is not important, the way how a question is asked or the choice the respondents have to choose from will affect the sample space (Ω). The choice that the respondents answered the questions affects directly the elements in Ω. So, it automatically will affect the probability of certain answer being chosen by the respondents.

Pretest

The above discussion shows the importance of conducting a pretest before administering any surveys. A pretest gives the opportunity to the researchers to reword or redesign the questionnaires in such away to achieve the objectives of the study in the most efficient way. Questions may also be deleted or added depending on its appropriateness. Most literatures in this area suggest that deleting or adding any particular questions should be done in line of internal reliability test. This test is measured by coefficient alpha as presented in Equation (1). Though there is no exact number for acceptable alpha coefficient, researchers tend to use a general rule. The closer the coefficient is to one the better it will be. As previously discussed, the closer the coefficient to one the better the surveys are because the questions measure what they are intended to measure. In both marketing research or psychology literatures, alpha is considered adequate if it equals to 0.85 or better. Deleting a question or an item from the set of questions in the surveys without lowering the coefficient is a good indicator that

the deleted question does not make important contribution to the whole study. The other important reason why a pretest needs to be done prior to actually conduct the surveys is to avoid expensive cost of repeating the study.

The other problem that worth mentioning in this discussion is the way one needs to work on the collected data. If one uses the questions as presented in experiment question#1, how does one need to key in the responses? Especially if respondents have chosen more than one answers. How do multiple answers need to be recorded or key-in in the data base so that it will be useful in the next step of the research process?

Discussions

After discussing so many encountered problems with experiment questions#1, the students made corrections on the choices that a respondent can choose by asking the following four set of possible revised questions with their respective choices.

Experiments question #2: Which of the following factors do you consider when you buy the bread?

- ☐ Price
- ☐ Softness / Freshness
- ☐ Nutrition content
- ☐ Packaging
- ☐ Convenience / Availability
- ☐ Price and softness
- ☐ Price and nutrition content
- ☐ Price and packaging
- ☐ Price and convenience / Availability

Experiment questions #3: Which of the following factors do you consider when you buy the bread?

- ☐ Price
- ☐ Softness / Freshness

☐ Nutrition content

☐ Packaging

☐ Convenience / Availability

☐ Price, softness/freshness and nutrition content

☐ Price, softness/freshness and packaging

☐ Price, softness/freshness and convenient

Experiment question #4: Which of the following factors do you consider when you buy the bread?

☐ Price

☐ Softness / Freshness

☐ Nutrition content

☐ Packaging

☐ Convenience / Availability

☐ Price and softness

☐ Price and nutrition content

☐ Price and packaging

☐ Price and convenience / Availability

☐ Price, softness/freshness and nutrition content

☐ Price, softness/freshness and packaging

☐ Price, softness/freshness and convenient

Experiment questions #5: Which of the following factors do you consider when you buy the bread?

☐ Price

☐ Softness / Freshness

☐ Nutrition content

☐ Packaging

☐ Convenience / Availability

☐ Price and softness

☐ Price and nutrition content

☐ Price and packaging

☐ Price and convenience / Availability

☐ Softness / Freshness and nutrition content

☐ Softness / Freshness and packaging

☐ Softness / Freshness and convenient/availability

☐ Nutrition content and packaging

☐ Nutrition content and convenient/availability

☐ Price, softness/freshness and nutrition content

☐ Price, softness/freshness and packaging

☐ Price, softness/freshness and convenient

☐ Price, softness/freshness, nutrition content and packaging

☐ Price, softness/freshness, nutrition content and availability

☐ Price, softness/freshness, nutrition content, packaging and availability

Students thought that by adding more choices as shown in experiment question #2, 3, 4 or 5, they might be able to capture the consumers' purchase decision. However, they certainly do not realize that even all possible combinations of choices are captured in experiment question #5, the students still having a big potential problem associated with the way the sample are drawn or the way the sampling was conducted. Experiment question #2 has mixed two alternative ways on how to draw the sample. The first five-question deals with drawing only one factor while the fifth through the ninth choices have asked the respondents to choose a combination of price and one of any other attributes. If the respondents are asked to choose only one factor than the combination of 5 and 1 $\left\{C_{r=1}^{n=5}\right\}$ will yield total elements of 5 in the sample space Ω. However, if the sample is drawn in a 2-tuple fashion, then one deals with the following combination $\left\{C_{r=2}^{n=5}\right\}$ which yields ten events in Ω. As one can see, different sampling techniques yield different Ωs which in turns will affect the probability of an outcome or a choice being picked.

After a long discussion on how to minimize the problems, students come with a great idea to ask the respondents the same question for each attribute and redesigned the choice with a Likert scale. The following questions were asked for each attribute:

Experiment questions #6: Price is the most important factor when I buy a loaf of bread

☐ Strongly disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Softness/ freshness is the most important factor when I buy a loaf of bread

☐ Strongly disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Nutrition is the most important factor when I buy a loaf of bread

☐ Strongly disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Packaging is the most important factor when I buy a loaf of bread

☐ Strongly disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Convenient/availability is the most important factor when I buy a loaf of bread

☐ Strongly disagree ☐ Disagree ☐ Agree ☐ Strongly Agree

Given the original question (experiment question #1), one can see the probability that a respondent will choose one particular choice is 0.20. However, if the questions are changed as shown in experiment question #2, then the probability that a respondent chooses a certain answer decreases to 0.111. Presenting the questions as shown in experiment question #2 is not correct because Ω has different elements. As the element changed, it also will affect both the expected value and the variance. The same mistakes

and violations of the statistical concept also found in experiment questions #3, #4 and # 5. This example demonstrates on how inappropriate survey design could alter the analysis or conclusion that one might have in a negative way.

On the other hand experiment questions #6 meets the criteria for a sample space theory and the criteria of appropriate wordings. In experiment question #6, one finds how helpful it is when an appropriate question is asked. Using questions as presented in experiment 1 through 5, one can not do the estimation because there is only one column matrix that can be constructed. With only one column matrix of observations no normal equation can be formed which prevents researchers of doing any estimation procedures. However, using experiment question #6, the researcher can find five columns made up of observations on each attribute which formed the attribute vector. Given these data he or she can transform the attributes observations into a new set of linear combinations as shown in Equation (2). The linear combination of the attribute accounts for the maximum variance can be used as a basis for a principal component analysis.

(2). $$Y_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + ... + a_{1p}F_p$$

Y_1 is called the first principle component, and if the coefficients are scaled such that $a_1'a_1 = 1$, then the variance of Y_1 is equal to the largest eigenvalue of Σ, the sample covariance matrix (Stevens, 1996, page 363). This is true because the coefficients of the principal component are the elements of the eigenvector corresponding to the largest eigenvalue. The second linear combination (Y_2) which is uncorrelated (Pearson correlation between Y_1 and Y_2 is zero) with the first component is searched such that it accounts for the next largest amount of variance in the system (after removing variance attributable to the first component). The coefficients of the second component are the

eigenvector associated with the largest eigenvalue of Σ and the sample variance of Y_2 equals to the second largest eigenvalue. Using this transformation approach enables the students to find the answers of their interests.

What Next

The Likert scale are commonly found and used in consumers' behavior study. One important note that needs to be addressed concerning Likert scale applied in surveys is that it generates a categorical data where the orders are important. When the order is important, one could apply either a multivariate or a categorical statistical data analysis in her or his approach depending on the research questions being studied. Agresti (1995) argues that categorical variables are variables for which the measurement scales consist of a set of categories. There are four types of categorical variables, but nominal or ordinal variables are commonly used in social science research. Nominal variables are those variables which have unordered scale. On the other hand, the ordinal data have a natural ordering such as the consumer's attitude or preference toward a product, a policy or a brand. The bread study as expressed in experiment question #6 is an example of a study where ordinal data are generated. When order is important then the methodology to analyze the data is design to take into account the importance of order in the estimation process. This quite different with the nominal data where the same results are obtained no matter in what order the categories are listed.

In the bread example, the purpose of the study is to find what factors, traits or attributes that affect consumers' purchase decision. So, there is no response variable needed to answer the research questions. Instead of using the usual prediction commonly found in most econometric approaches, one might be able to use either principal

component or factor analysis or SEM (Structural Equation Model) or also known as LISTREL (Linear Structural Relationships) to answer the research questions. The principal component and factor analysis are the most common approaches within multivariate statistical analyses to find the answer of consumer's attitude such as the one shown in the bread study. But one might need also to consider of using the SEM with latent variable or latent factor which are believed to be important and can add explanation of buyers' purchasing attitudes (Bollen, 1989). He continued arguing that SEM can be viewed as a regression analysis approach with less restrictive assumptions that allows ME to occur in both explanatory as well as in the response variable. SEM also consists of factor analyses that permit direct or indirect effect between factors. Therefore, according to Bollen, SEM encompasses and extends regression, econometrics, and factor analysis procedures.

Experiment Results

The results of principal component analyses with varimax rotation method estimated on collected observations (there are total 47 of them) showed that there are three factors consist of combination of two different attributes which considered important by the bread buyers (Table 1).

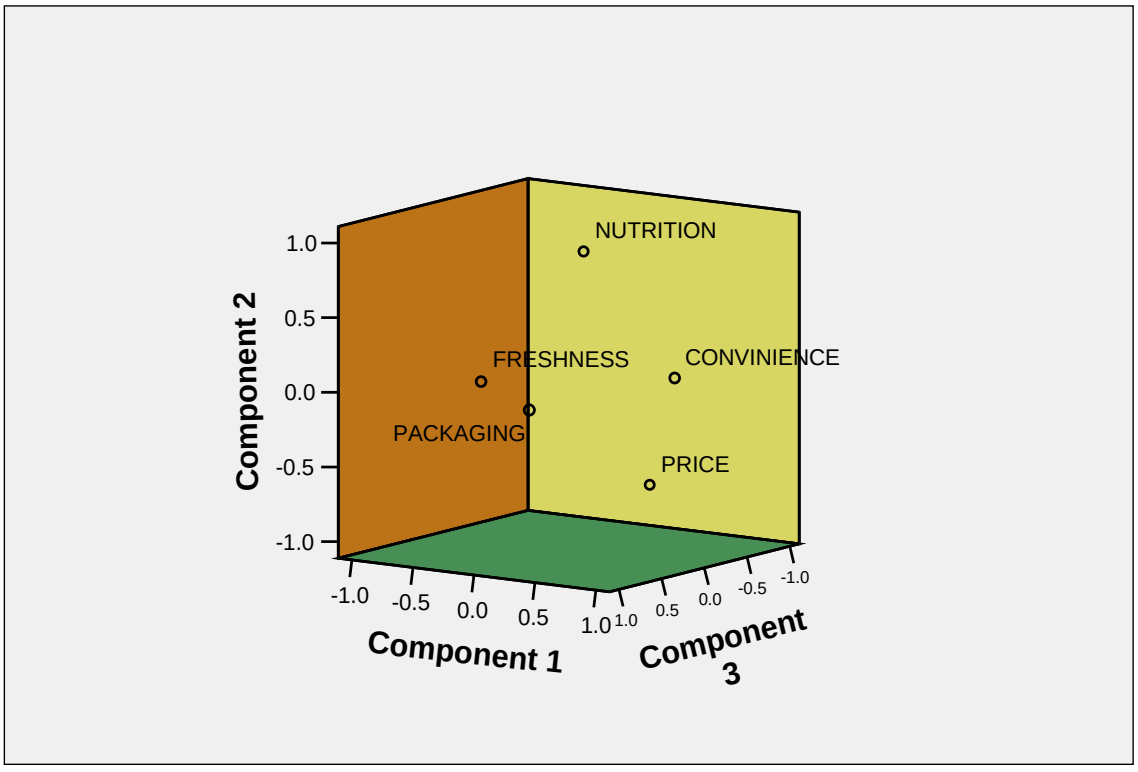
Table 1 - Rotated Component Matrix

	Component		
	1	2	3
PRICE	.618	-.614	-.066
FRESHNESS	-.235	.100	.691
NUTRITION	.146	.916	.035
PACKAGING	.217	-.033	.772
CONVINIENCE	.879	.138	.015

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Though the analysis does not have a significant goodness of fit as measured by Bartlett's Chi-square statistics, it demonstrates that asking the original question as shown in experiment question #1 causes a bias and misleading results. The insignificant of the goodness of fit measure is due to a small sample size. This flaw is also confirmed by KMO (Kaiser-Meyer-Olkin) statistics on sampling adequacy. However, the sole purpose of this paper is just to show that inappropriate wordings or making vague choices need to be avoided. Figure 1 shows the components plot in the rotated space.

Figure 1 - Component Plot in Rotated Space



—

Conclusion

A simple example on how inappropriate asked questions in surveys could jeopardize the results of a study has been presented. Many steps need to be done appropriately to design and redesign surveys questions if they are chosen to be a tool of data collection. In recent years, more researchers in Agricultural Economics profession have used surveys to collect consumers' opinions, attitudes on certain product attributes or government policies. Given its raising popularity, researchers also need to be aware of limitations that surveys might have. Certain steps need to be done suitably to ensure that the surveys achieved what they are supposed to and designed for.

References

Agresti, Allan., An Introduction to cathegorical Data Analysis. Wiley Series in Probability Statistics, 1996.

Armstrong, G. D. The Interclass Correlation as a Measure of Interrater Reliability of Subjective Judgments. Nursing Research 30(1981): page 314-320.

Bentler, P.M and Yutaka Kano. On the Equivalence of Factors and Components. Journal of Multivariate Behavioral Research 25(1990): page 67-74.

Bollen, Kenneth A. Stuctural Equations With Latent Variable. Wiley Series in Probability and Mathematical Statistics. John Wiley & Son. 1989.

Cohen, Jacob. The Cost of Dichotomization. Applied of Psychology Measurement 7(summer 1983): page 249-253.

Cortina J.M., What is Coefficient Alpha? An Examination of Theory and Application Journal of Applied Psychology 78(1993): page 98-104.

Crocker, Linda and James Algina. Introduction to Classical & Modern Test Theory. Harcourt Brace Jovanovich College Publishers. 1986.

Cronbach, J. Lee., Coefficient Alpha and the Internal Structure of Tests. Psychometrika 6(1951): page 297-334.

DeVellis, Robert F., Scale Development: Applied Social Research Methods Series Volume 26. Sage Publications. 1991.

Dillman Don A. Mail and Internet Surveys: The Tailored Design Method. JohnWiley &

Sons. Second Edition, 2000.

Djunaidi, Harjanto., "The Attitude of MTSU Students and Employees toward MTSU Milk". Southern Agricultural Economics Association Annual Meetings. Little Rock. 2005

Djunaidi, Harjanto., Attitude Assessment of the University of Arkansas Students and Employees toward Food and Medicine Taxes. Research was conducted in Measurement and Evaluation Class. Department of Education Foundation and Statistics, the University of Arkansas, Fall 2002.

Djunaidi, Harjanto., Finding Common Communalities which Affect Graduation Rates in the U.S. Universities: A Principle Component Analysis Approach. Research was conducted in Multivariate II Class, Department of Statistics, University of Arkansas, Fall 2001.

Gorsuch, Richard L. Common Factor Analysis versus Component Analysis: Some Well and Little Known facts. Journal of Multivariate Behavioral Research 25(1990); Page 33-39.

Greene., William H., Econometric Analysis., Third Edition. Prentice Hall1997.

Hair, Joseph F, JR, Rolph E. Anderson, Ronald L. Tatham and William C. Black, Multivariate Data Analysis. Prentice Hall. Fifth Edition, 1998.

Johnson, Richard A and Dean W. Wichern. Applied Multivariate Statistical Analysis. Prentice Hall. 1998.

Kelloway, E. Kevin and Julian Barling. Item Content versus Item Wording: Disentagling Role Conflict and Role Ambiguity. Journal of Applied Psychology 6(1990): page 738-742.

Maddala, G.S., Introduction to Econometrics. MacMillan Publishing Company. Second Edition. 1992.

Mood, Graybill and Boes., Introduction to the Theory of Statistics., McGraw-Hill Series Company in Probability and Statistics. Third Edition. 1963.

Robertson, Michael T, and Eric Sundstrom. Questionaire Design, Return Rates, and Response Favorableness in an Employee Attitude Questioners. Journal of Applied Psychology 3(1990): page 354-357.

Steven, James. Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum Associates Publishers. Third edition. 1996.

Velicer, Wayne F, Douglas N. Jackson. Component Analysis versus Factor Analysis: Some Issues in Selecting an Appropriate Procedure. Journal of Multivariate Behavioral Research 25(1990): page 1-28.