AN ABSTRACT OF THE THESIS OF

Diego Andrade for the degree of Master of Science in Agricultural and Resource Economics presented on November 7, 2003.

Title: The Impact of Health and Environmental Motives and Economic Factors on the Choice of Organic Produce.

Abstract approved: Signature redacted for privacy.

Catherine A. Durham

This research examines the factors that influence consumer choice of organic over conventional fresh produce. The emphasis of this study is to determine the relative importance of environmental and health motives and economic factors in explaining this choice. A survey was designed to collect data about consumer preferences and purchasing behavior in three locations in the Portland metropolitan area: a conventional supermarket, a farmers' market, and a cooperative. These locations were selected to increase the diversity of the sample population with respect to organic purchasing. Data was collected using tablet personal computers. Analysis was undertaken using the binomial logit model. All models were successful in explaining the choice of organic fresh produce in the sample. The results show that both environmental and health motives are important in explaining consumer choice of organic produce. Age and price sensitivity were also found significant in explaining consumer preference and purchasing of organic versus conventional produce while income was significant in differentiating buying organic from the preference for organic.

The Impact of Health and Environmental Motives and Economic Factors on the Choice of Organic Produce

by

Diego Andrade

A THESIS

Submitted to

Oregon State University

In partial fulfillment of
the requirements for the
degree of

Master of Science

Completed November 7, 2003
Commencement June 2004

Master of Science thesis of Diego Andrade presented on November 7, 2003

APPROVED:

Signature redacted for privacy.

Major Professor, representing Agricultural and Resource Economics

Signature redacted for privacy.

Chair of Department of Agricultural and Resource Economics

Signature redacted for privacy.

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Signature redacted for privacy.

Diego Andrade, Author

ACKNOWLEDGEMENTS

These pages mark the culmination of another chapter in my life. An intense, rich and joyful time well spent between the faculty and colleagues of the AREC department, from whom I'm taking invaluable knowledge and unforgettable memories.

I want to specially acknowledge the guidance and time provided to me by Dr. Catherine Durham, who was not only my advisor in the academia but a mentor in life. From Cathy I am taking the example of a dedicated professional and a generous teacher, who shared her knowledge and time with me.

I would also like to thank the members of my committee Aaron Johnson, Larry Lev and John Selker for their valuable comments and suggestions, which helped me improve not only this study but my professional skills. A special acknowledgement for the members of the Food Innovation Center who helped me both in human and capital resources in completing this study.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure | Page

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF APPENDICES

LIST OF APPENDIX FIGURES

LIST OF APPENDIX TABLES

DEDICATION

To my mother, the source of my energy
and the meaning of my life.

The Impact of Health and Environmental Motives and Economic Factors on the Choice of Organic Produce

1. Introduction

1.1 Introduction

Modern organic agriculture began in the 1940's, emphasizing that health of plants, soil, livestock and people are interrelated. Organic farming practices were developed to work in harmony with nature using natural inputs produced on farm. During the 1960's and 1970's the organic scope was broaden by the inclusion of the relationship between agriculture and resource conservation with emphasis on the use of nonrenewable resources. By 1980, public concern regarding pollution caused by traditional farming practices was awakened, bringing a new wave of buyers for organic products.

Recently, organic farming has been one of the fastest growing segments in U.S. Agriculture. Certified organic farming grew from 0.9 million acres in 1992 to 2.3 million in 2001. U.S. retail sales of organic foods have been growing at 20-24% during the last 12 years, with an estimated level of $11 billion in 2002. For 2003 it is estimated that organic sales with be around $13 billion with a projection of $19 billion for 2010. Organic is a growing segment, but it still has small share of the total food sales market. According to the United States Department of Agriculture (USDA) it is estimated that for 2003 the share of organic food and beverages sales to be 2.5% of the total food retail sales. Organic products are now available in over 20,000 stores. They are sold in natural food stores, in farmer's markets and in 73% of conventional grocery stores.

In 1990 the Organic Food Protection Act (OFPA) was passed initiating new regulations including the requirement for producers to follow an organic plan accepted by an accredited certifying agent. The OFPA mandated the creation of the National Organic Standards Board (NOSB) to advise the Secretary of Agriculture on organic standards. The OFPA is enforced by the National Organic Program (NOP) part of Agricultural Marketing Service of the U.S. Department of Agriculture (USDA). In October 21, 2002 the final rules of the national standards went into effect. Food products that are labeled organic must meet these standards, regardless of whether the product is grown in the United States, or imported from other countries.

According to the NOSB,

> "Organic agriculture is an ecological production management system that promotes and enhances biodiversity, biological cycles and soil biological activity. It is based on minimal use of off-farm inputs and on management practices that restore, maintain and enhance ecological harmony".

Organic farming is a particular production process that differs from conventional farming in the use and management of the crop by using natural means, avoiding synthetic chemicals and adopting practices that avoid damage to soil, water and other resources.

Many studies have been conducted by researchers in the public and private sectors on the buying habits and demographics of consumers of organic and eco-friendly foods. Specific results have varied depending on the type of survey, sample size, and geographic coverage. However, a few general themes have emerged. One outcome is that consumers prefer organically produced food because of perceived health attributes and concerns about pesticide residues, the environment (such as soil, water quality and wildlife habitat), and farm worker safety.

Some studies indicate that consumers perceive "health attributes" embedded in organic products as one of the major criteria for buying them besides environmental friendliness. The reduced use of pesticides and synthetic chemicals in the production process has given consumers the perception that organic produce is healthier and less risky. Moreover, studies have found that organic products are perceived as more "natural" than their conventional counterparts.

The NOP labeling standards don't allow claims that "organically produced food is safer or more nutritious than conventionally produced food". However, as noted above a few studies have linked the health conscious consumer to organic buying. Environmental protection and healthfulness of organic products are not competing attributes, and they may be complementary. The principle goal of this research is to examine more fully how the perceptions and behavior of consumers towards these two issues influence the choice of organic products.

In the next 8 years, organic sales in the U.S. are expected to increase about 20-25% annually, to about 19 billion dollars (See Figure 1.1). In these circumstances the knowledge and beliefs of the organic consumer is extremely important for the food retail industry as well as for the production stage of the agricultural value chain. Moreover, given the potential environmental benefits of organic production, and public interest, it is important for policy consideration. Identifying the main attributes that influence the choice of organic products over conventional produce will help to obtain a better understanding of the organic consumer and develop better marketing strategies.

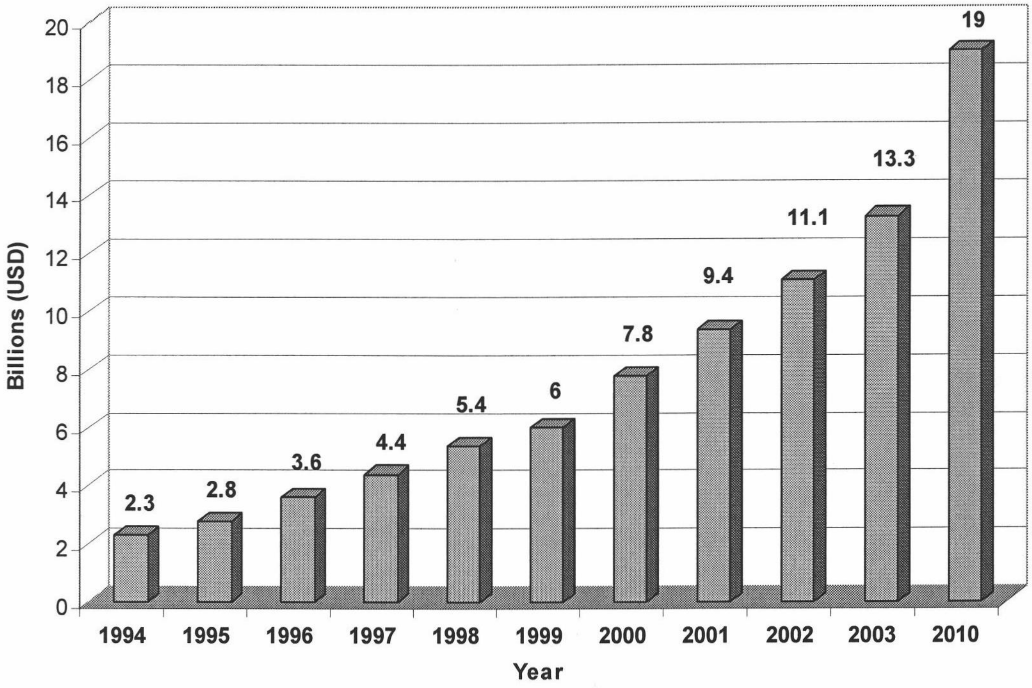


Figure 1.1. The Organic Food Industry Growth in the US

Source: Myers, S., and Rorie, S.
Note:

- *2002 and 2003 are forecasts form the Organic Trade Association*
- *2010 forecast by the Canadian Market Research Center, April 2001*
- *Organic Food Sales in 1980 in the U.S. were $78 million*

1.2 Objectives

The overall objective of this study is to determine the most important variables that determine the choice of organic over conventional fresh produce. The research will focus on the role of consumer's behavior towards the environment and health in explaining the choice of organic. This study will use fresh produce as its object.

Specific objectives of this study are to:

- To determine the relative importance of environmental and health motives in explaining the choice of organic produce.
- To make use of "behavioral questions" instead of "attitudinal questions" to elicit consumer's perception toward the environment and their health.
- To compare the most important variables that influence preference for organic produce with those influencing the purchasing decisions.

2. Literature Review

During the last two decades many researchers have examined consumer demand for organic products. The dynamic of the organic market created the need for information about the organic consumer, their demographics, shopping styles and primary interests. Researchers in the academic and private consulting sector have used a wide variety of methodologies and regional scopes that yielded a considerable amount of literature. The findings and methods for some of the most important studies, both academic and industry, about consumer demand for organic products are presented in section 2.1. Relevant studies about the effects of health and environmental motives on consumer demand for food products are presented in section 2.2. A table summary of the results of the principal studies on this area is presented in Appendix 2.

2.1 Studies on Consumer Demand for Organic Products

Thompson (1998) lists the most important studies about the characteristics and demographics of green consumers before 1997. Those studies are Baker and Crosbie (1993), Parkwood Research Associates (1994), The Hartman Report (1996), Food Marketing Institute/Prevention Magazine (1997) and The Packer (1997). The "Hartman Report" was the first comprehensive effort to define the market from a consumer perspective for sustainable agricultural products. The Hartman Report identified 6 segments in the American market involving their attitude towards the environment. The majority of the American public was found to be too "overwhelmed" (30%) with their personal economic situation to worry much about the environment. Only 7% of the consumers were assessed to be "*true naturals,*" meaning that they have a true commitment to save the environment, overcome price barriers and buy environmentally sound products.

In 1996, Huang used a bivariate probit model to study consumer preferences in Georgia (U.S.) for organically grown produce and to determine the importance of socio-demographic characteristics that may affect their willingness to purchase organic produce if it has sensory defects. Huang found that people who ranked pesticide residues on food as one of their top food concerns are more likely to prefer organically grown produce. Moreover, the study found that consumers who believe that fresh produce should be tested and certified residue free, and are nutritionally conscious, prefer organic to conventional produce. Consumers who consider low price as an important attribute are less likely to indicate a preference for organically grown produce while appearance of produce was not found significant.

Thompson (1998) reviewed the principal studies of the 1990s and concluded that "Studies of consumer demand for organic products have relied almost exclusively on self-reporting of purchasing behavior and attitudes as elicited through questionnaires and interviews; direct observation of consumer behavior at retail markets is almost nonexistent". He summarizes the findings of the most important published studies depending on the principal demographic variables: age, income, education and gender. Most of the studies about consumer demand have emphasized organic demographic characteristics to create a profile for the organic consumer. Thompson found that national studies[1] generally suggest that higher income households are more likely to purchase organic products but there are also studies that have identified low income segments that show high preferences for organic. This result suggests that higher income does not necessarily lead to higher preferences. He also points out that consumer's choice of shopping location is influenced by their disposition to buy organic grown goods. According to Thompson, there is some statistical evidence that the effect of age on organic purchases is noticeable for certain segments of the population, specifically for people over forty years old and for young consumers under thirty. He also points out that there is some evidence that suggests gender and marital status are important factors in influencing the choice of organic. He also concludes that education attainment has been found important in some studies while household size has been the least investigated demographic characteristic with mixed results.

Thompson and Kidwell (1998) designed a bivariate probit model to explain the choice of organic fresh produce by measuring choices of organic and conventional products in two retail outlets (one specialty store and one cooperative) in Tucson, Arizona. In addition to demographic variables, they included cosmetic defects and store choice in the model. They found that the larger the difference in the number of the cosmetic defects between organic and conventional produce, the less likely were consumers to purchase organic produce. However, the effect was small. The effects of price differences between organic and conventional were statistically

[1] Thompson refers as national studies to the ones carried by Parkwood Associates (1994), The Hartman Group (1996-1997), The Packer (1998).

significant for the specialty store with a relatively small marginal effect (-0.0116) per dollar. They found that demographic variables like age and gender were not significant, while households with more children under the age of 18 are more likely to buy organic (0.10). The effects of education on the decision to purchase organic were mixed. Having a college degree had no significant effect on choosing organic, while having a graduate or professional degree apparently decreased the probability of choosing organic (-0.159). They concluded in the study that the choice of organic fresh produce seems to be closely linked to the choice of store in which consumers usually shop.

In a 1999 study, Glaser and Thompson examined and compared the demand for organic and conventional frozen vegetables using national supermarket scanner data. Price and expenditure elasticities were estimated using an AIDS model. This study is one of the few that used scanner data instead of consumer stated preferences. They estimated four separate demand equations for broccoli, green beans, green peas and sweet corn. Own price elasticities for organic vegetables were found much larger than for conventional products ranging from -1.630 to -2.26. Cross price elasticities showed large standard errors, and there appears to be a tendency toward asymmetry in cross-price response. Expenditure elasticities found for organic frozen vegetables seem to suggest that increases in real income may not generate huge gains in organic market shares.

Loureiro and Hine (2001) used a multiple bounded probit model to construct three individual parametric estimations of willingness to pay a premium for value-added potatoes labeled as “Colorado-grown”, “Organic” and “GMO-free”. The study used a survey at the point of purchase in a sample of Colorado supermarkets retrieving socio-demographic information. They found that consumers were willing to pay a higher price premium for "Colorado grown" (5.52 cents /lb) than for organic (3.13 cents/lb) and "GMO-free" (0.16 cents/lb) potatoes. It was found that "upper-class" consumers (high income and with graduate education), and those concerned with freshness are willing to pay more for organic potatoes. Their results also showed a negative relationship between presence of children in the household and paying a premium, while gender was not found significant in any of the three equations.

Recently, Wang and Sun (2003) conducted a Conjoint Analysis to examine consumer preferences and evaluation of important attributes of organic apples and milk. Four attributes were chosen for the analysis: production method, certification, origin and price. Results indicate that consumers who purchased organic food preferred the following level of each of the four attributes: organically grown, produced in Vermont, certified by NOFA and priced at 0.99$. For the non-organic buyers, the attribute levels found were conventionally grown, produced in Vermont, certified by USDA and priced of $0.99. For organic buyers the most important attribute is location with a relative importance of 31.67% while for non-organic buyers price is the most important attribute with 49.27% of relative importance. The authors report that when consumers where asked to answer the principal reasons for buying organic, 69% reported because it was healthier, 55% to help small farmers and 54% because it was better for the environment.

Most of the studies on consumer demand for organic food products have been done on consumer's preferences; few studies have made use of observed retail data as noted in this section. Moreover, consumer's preferences have been approached using mainly discrete choice models. Also, the results regarding demographic factor's relationship to consumer's preferences for organic food have varied from study to study. No definite conclusions can be made about the relationship between organic preferences and a number of demographic factors.

2.2 Studies on Environmental and Health Motives in Demand for Food Products

This section reviews the results obtained from studies that have analyzed the effect of environmental and health motives in the demand for food products. Since one of the objectives of this study is to analyze the effect of environmental and health factors influencing organic preferences and purchases of organic produce, a brief overview of the literature on these factors is needed. Most of the summaries presented in this section come from academic research studies.

Eom (1994) developed a random utility model integrating consumer stated preferences for safer produce with their risk perceptions in response to scientific information about pesticide residues and subjective attitudes towards pesticide residues in North Carolina. He used a binomial Probit framework modeling the choice of conventional produce tested for pesticide residues incorporating a technical measure of risk taken from National Academy of Science (NAS) and Environmental

Protection Agency (EPA) studies and a subjective measure of risk measure by a scale. It was found that the technical measure of risk did not affect purchase intentions but the scalar variable measuring subjective health risk of consuming traditional produce did have a significant, positive effect on choosing the tested produce. It was also found that variables measuring price differences and education were inversely related to the choice of safer produce. Variables measuring income, age, and number of children were not significant. Finally consumers were willing to pay considerably more for safer produce, with price increments of 85-90% relative to the base price of non- tested produce.

Byrne, Bacon and Toensmeyer (1994) used a logit model to determine the effect of socio-demographic characteristics on levels of pesticide residue concerns and determine the impact of these effects on shopping location behavior. They collected data using a mail survey in Delaware where they included a set of attitudinal questions about the concerns of the effects of pesticide residues on the environment and health. The results showed that those giving high importance to nutrition and environmental effects have a significantly higher probability to be concerned with pesticide residues. The marginal effects for these two variables were 0.13 and 0.12 respectively. Moreover, consumers who disagree that risks associated with the food supply have been exaggerated are 11% more likely to be concerned with pesticide residues. Demographic variables including age, gender, education, income, being single and having children were found not to be significant. Concern regarding pesticide residues, increases probability that consumers would shop at a supermarket that offered pesticide residue free produce, even at higher prices.

Estes and Smith (1996) used a hedonic framework to examine price, appearance, and health risk considerations made by shoppers in Arizona. They examined how people evaluate tradeoffs among price, health risk and other dimensions of quality when they make fresh produce purchase decisions. They observed retail prices for organic apples were on average 118% higher than conventional apples. Statistical analysis suggested that consumers evaluated size, weight, defects, organic labeling, variety and package size in their purchasing decisions. Sensory defects seemed to have no significant effect on retail prices for organic or conventionally grown produce. They concluded that the primary reason for higher organic prices is unrelated to appearance features but to the perception that organic conveys a lower risk of exposure to pesticide residues. They also concluded that the higher prices paid by consumers is explained perhaps by the fact that buyers were purchasing additional food safety by eliminating uncertainty about pesticide residues.

Blend and van Ravenswaay (1999) studied consumer demand for ecolabeled apples examining consumer purchasing intentions, the effect of environmental claims and the role of concerns about the environment and food safety. They used both a Cragg Double-Hurdle and a Tobin Model using a national sample of consumers. It was found that consumers who believe ecolabeled products might improve the environment are more likely to buy ecolabeled apples with a marginal effect of 0.55. Moreover, the variable measuring food safety concern (1= buy to improve health/safety, 0=otherwise) was not significant, though it had a positive sign. Purchase probability for labeled apples is significantly affected by own price, unlabeled apple price, familiarity with IPM, gender, education and month of purchase. Comprehensiveness of environmental claims and the amount of environmental proof were not found to affect either the purchase probability or the quantity purchase of ecolabeled apples.

Govindasamy and Italia (1999) used a Logit framework to empirically evaluate which demographic characteristics cause consumers to be more likely to pay a premium for organically grown produce in New Jersey. Besides the traditional demographic variables they included one variable measuring the respondent's perception if health risk of synthetic pesticides and another one measuring if individuals believe pesticides have negative effects on the environment. It was found that smaller households and higher earning households would be more likely to exhibit a higher willingness to pay for organic produce. Also younger households in which females are the primary shopper also appear to be more likely to pay a 10% premium for organic produce. Furthermore a house with more knowledge about alternative agriculture such as IPM, and which usually or always buys organic are more likely to pay the premium. The findings also suggest that education and willingness to pay for pesticide reduced produce are inversely related. Neither the variable measuring the perception of pesticide's health risk or the one measuring the negative effects of pesticides on the environment were significant and exhibit low marginal effects (0.02 and 0.07 respectively).

Dimitri and Richman (2000), in reviewing the Hartman Report titled "The Evolving Organic Marketplace" (1997), stated that the five top criteria for consumers to purchase organic were: healthfulness (80%), availability in regular supermarkets

(69%), environmental friendliness (67%), price (64%) and convenience of preparation (53%). They also review a report commissioned by the Food Alliance in which 600 organic consumers from Portland, OR, were asked which are the most important food qualities. The top five qualities rated as "extremely important" were: absence of synthetic pesticides (77%), absence of synthetic herbicides (77%), absence of e. coli and other harmful bacteria (75%), absence of artificial ingredients or preservatives (61%) and absence of synthetic fertilizers (59%). Quoting Dimitri and Green (2002), "health and environmental issues are of paramount importance to consumers interested in organic foods".

Lourerio, McCluskey and Mittelhammer (2001) used a multinomial logit approach to identify the sociodemographic characteristics affecting the choice among organic, eco-labeled and regular apples in Portland, OR. Besides demographic information, they also collected respondent's attitudes towards the environment and food safety. They found that the probability of choosing organic apples increases when consumers have children under 18 and when they have strong attitudes towards the environment and food safety. They found that the marginal effect of the variable measuring environmental attitudes was positive, significant (0.36) and higher than the Food Safety variable (0.16). Family size has a negative effect while race, age and education were not statistically significant. The study concluded that eco-labeled and organic apples appeal to consumers with strong attitudes towards food safety and environmental quality, with the eco-labeled being the intermediate choice between organic and regular apples.

Govindasamy, Italia and Adelaja (2001), followed a logistic approach to evaluate the demographic characteristics that influence consumers to pay a premium for integrated pest management (IPM)[2] produce. Being male and being over 65 years of age have negative effects on the probability of paying a 10% premium for IPM. Higher income people who usually purchase organic produce have a positive and significant relationship with the probability of paying a premium for IPM produce. It was found also that consumers with high risk aversions (people who believed that the use of synthetic pesticide posed a very serious health risk) towards pesticides were 16% more likely to pay a premium.

Dimitri and Greene (2002) reviewed the latest industry and academic studies about consumption characteristics of the U.S. Organic Sector. They listed the main reasons for consumers to purchase organic according to the "Hartman Group" Report in 2000. In first place are health and nutrition (66%), followed by taste (38%), environment (26%), and availability (16%). Dimitri and Green also review the results from the "Walnut Acre Survey" in which 63% of the respondents were found to believe that organic food and beverages were more healthful than their conventional counterparts.

Given the results of the studies presented in this section, one important observation can be made. With the exception of Govindasamy and Italia (1999), the rest of the studies found environmental concern to be important and correlated with the dependent variable analyzed in each of the studies. In the case of health related concerns, evidence of its importance is mixed. Moreover, the studies reviewed in this

[2] IPM uses a natural approach to pest control minimizing the dependence on synthetic chemicals. IPM falls between conventional and organic agriculture according to the literature.

section have used mostly attitudinal questions to elicit respondents about environmental and health concerns. They have used questionnaires asking consumers about their health risk perceptions, beliefs and opinions about food safety and environmental concerns for each particular case.

2.3 Current Approach

The current study covers new ground in a number of respects. First, although there are a number of studies available on consumer demand for organic products, this is the first one to focus on comparing the environmental and health factors influencing the choice of organic food.

Second, this study uses both environmental and health variables principally obtained by the use of behavioral questions rather than single scale or dichotomous variables of perceptions and attitudes in the choice of organic. Behavioral questions seek factual information with respect to what people do, how often, how they use, etc., regarding a particular issue while attitudinal questions seek information about people's beliefs, opinions or intentions. A technique called factor analysis is used to reduced and condense the information obtained from these questions.

Finally, this is the first study to examine whether the majority of the respondent's produce purchases are organic. Most of the studies on consumer's preferences for organic produce have relied on purchase intentions. This study includes in the analysis a variable that measures the relative participation of organic fresh produce purchases. Both preferences and purchases for organic produce are compared and analyzed by using a discrete choice model derived from the random utility theoretical framework.

3. Theoretical and Empirical Model

Research on consumer preferences has intensively used categorical limited dependent variables to model consumer's choice among a set of alternatives. In this type of model the dependent variable is categorical and discrete, taking values that represent a specific choice. There are four basic approaches of deriving models for categorical limited dependent variables: as linear probability models, as latent dependent variables, as non-linear probability models and as discrete choice models.

One of the objectives of this study is to model the choice of organic fresh produce in order to elicit the factors that influence the consumer's choice of organic over conventional produce. The discrete choice model is well suited for the task since it is based on the Random utility theory that models the choice of alternatives given the relative level of utility that each alternative yields to the consumer.

The next two sections of this chapter present the theoretical model on which discrete choice models are based. The microeconomic principles of choice demand theory and random utility theory are discussed. The third section covers the general empirical model and econometric specification of discrete choice models.

3.1 Choice and Demand Theory

In the next paragraphs the microeconomic theory of choice and demand, which is based on the economic concept of consumer preferences, and then random utility theory is reviewed. In the review of choice and demand theory emphasis is placed on the choice axioms of rational behavior, the utility maximization problem faced by the consumer and the derivation and properties of indirect utility function because of its relevance to discrete choice models.

This section relies on Microeconomic Theory by Nicholson (2002), Microeconomic Theory by Mas-Colell, Whinston and Green (1995), and Microeconomic Analysis by Varian (1992).

Consider a set of possible and mutually exclusive alternatives from which an individual must choose denoted by X. Traditional microeconomic theory characterized the individual by the decision maker's tastes which are summarized in a preference relation in which the consumer chooses the alternatives contained in X. This preference relation is denoted by "$\succsim$", which allows the comparison of pairs of alternatives $x, y \in X$. The relation $x \succsim y$ is read as "x is at least as good as y". The symbol $\succ$ denotes strict preference relationship; $x \succ y$ is read "x is preferred to y". While the symbol "$\sim$" represents indifference; $x \sim y$ is read "x is indifferent to y". For preference it should be understood that when an individual reports that x is preferred to y, it is taken to mean that all things considered, he or she feels better off under situation x than under situation y.

Microeconomic theory is based on the assumption that consumers are rational. The hypothesis that consumer behavior is rational is embodied in two basic properties of the preference relation which follow:

(i) Completeness: For all $x, y \in X$, individuals can specify $x \succsim y$ or $y \succsim x$ or both. In other words if x and y are two situations an individual is faced with, he or she can always specify one of the following three possibilities:

a. $x \succ y$: x is preferred to y

b. $y \succ x$: y is preferred to x, or

c. x ~ y: x and y are equally attractive

(ii) Transitivity: For all $x, y, z \in X$, if individuals specify $x \succsim y$ and $y \succsim z$, then $x \succsim z$. If x is preferred to y and y is preferred to z, then the individual must also report that x is preferred to z.

There are two other properties that are also reported by the economic literature:

(iii) Reflexivity: For all $x, y \in X$, x $\succsim$x, and,

(iv) Continuity: For all $x, y \in X$, the sets {x: $x \succsim$y } and {x: y $\succsim$x } are closed sets. Which suggests that if an individual reports x is preferred to y, then situations suitably close to x must be also be preferred to y.

Given the properties of completeness, transitivity, reflexivity and continuity, individuals are able to rank in order all possible alternatives from the least desirable to the most. The preference ordering can be represented by a continuous "utility function." Denote utility function as u: X → R such that $x \succ y$ if and only if u(x)>u(y).

Utility functions are ordinal in the sense that only the ranking of alternatives matters and are invariant for any strictly increasing transformation. If u(x) represents some preferences and f: R→R is a monotonic function then $f(u(x))$ will represent exactly the same preferences since $f(u(x)) \geq f(u(y))$ if and only if $u(x) > u(y)$.

In demand theory the decision maker is the consumer who chooses an alternative (good) from the choice set (consumption bundle). It is assumed that the consumer's rankings of these goods can be represented by a utility function of the form:

(1) $U=U(X_1, X_2, \ldots.., X_c)$

Where X_i (i= 1,2…., c) refers to the quantities of the goods that might be chosen from a bundle of c goods.

Given the consumption bundle, consumers will choose the quantities of the goods that yield the higher utility constrained by the amount of income the consumer has to spend (budget constraint). Indeed, the consumer faces a utility maximization problem of this form:

(2) Max $U=U(X_1, X_2, \ldots.., X_c)$

$$\text{s.t. } I = \sum_{i=1}^{c} P_i X_i$$

where P_i is the price of the good i and I represents income.

This utility maximization problem can be solved using a Lagrange Multiplier method:

(3) Max $L = U(X_1, X_2, \ldots.., X_n) + \lambda\, (I - \sum_{i=1}^{c} P_i X_i)$

Taking the partial derivatives of L with respect to X_i and setting them equal to zero yields c+1 equations necessary for a maximum:

$$
(4)\quad
\begin{aligned}
&\frac{\partial L}{\partial X_1} = \frac{\partial U}{\partial X_1} - \lambda P_1 = 0 \\
&\frac{\partial L}{\partial X_2} = \frac{\partial U}{\partial X_2} - \lambda P_2 = 0 \\
&\vdots \\
&\frac{\partial L}{\partial X_c} = \frac{\partial U}{\partial X_c} - \lambda P_c = 0 \\
&\frac{\partial L}{\partial \lambda} = I - \sum_{i=1}^{c} P_i X_i \; = 0 \\
&\qquad\qquad\qquad\qquad 1
\end{aligned}
$$

If the utility function is continuous and the constraint set is closed, the optimal quantities of the goods X_i (i=1,2,...c)and for λ can be obtained. The first order condition for a maximum is:

$$(5)\ \lambda = \frac{MU_{x_1}}{P_1} = \frac{MU_{x_2}}{P_2} = \cdots \frac{MU_{x_n}}{P_c}$$

Where MU_{xi} is the marginal utility of good *i* defined as $\frac{\partial U}{\partial X_i}$. This equation say that each good purchased should yield the same marginal utility per unit of currency spent on that good. The term λ is regarded as the marginal utility of income. The second-order condition requires that the Hessian matrix of the utility function is negative semidefinite which implies that the upper contour set is convex in the neighborhood of the optimal solution.

From (4) the optimal values of X_i $\forall\, i$ can be solved for, which will depend on the prices of all goods and on the individual's income. The optimal value of X_i, denoted as X_i^*, is known as the Marshallian Demand of the i^{th} good.

$$
(6)\quad
\begin{aligned}
X_1^* &= X_1(P_1, P_2, \ldots, P_c, I)\\
X_2^* &= X_2(P_1, P_2, \ldots, P_c, I)\\
&\vdots\\
X_c^* &= X_c(P_1, P_2, \ldots, P_c, I)
\end{aligned}
$$

Substituting the optimal values of the quantities of the goods into the consumer's utility function the "indirect utility function" is obtained showing the utility that the consumer would receive at the chosen quantities obtained from maximizing utility subject to the budget constraint. Substituting (6) into (2) yields:

(7) Maximum utility = $U^* = U(X_1^*, X_2^*, \ldots, X_c^*)$

$$= U^* = V(P_1, P_2, \ldots, P_c, I)$$

This is now as Direct Utility function $U(X_1, X_2, \ldots, X_c)$ that gives the utility that consumer obtains at given quantities of each good and a Indirect Utility function $V(P_1, P_2, \ldots, P_c, I)$ which gives the utility that the consumer obtains at given prices and income once he has chosen the optimal quantities of the goods. Varian (1995) shows that a consumer's preferences can be equivalently represented by either the direct utility or the indirect utility function. Using matrix algebra notation, let "p" be the vector of prices, let *v* represent the indirect utility function, and *m* be income. According to Varian (1995) the properties of the indirect utility function are:

1) $v(p,m)$ is nonincreasing in p; that is, if $p` \geq p$, $v(p`,m) \leq v(p`,m)$ and nondecreasing in m.

2) $v(p,m)$ is homogeneous of degree 0 in (p,m).

3) $v(p,m)$ is quasiconvex in p; that is $\{p: v(p,m) \leq k\}$ is a convex set for all k.

4) $v(p,m)$ is continuous at all $p > 0$, $m > 0$.

3.2 Random Utility Theory

As seen in the last section, having a set of alternatives defined on a continuous space allows the use of calculus to derive demand functions. However, if the choice set of alternatives are discrete, and the consumption of one or more commodities can be zero, the maximization problem will yield a corner solution where the first-order conditions for an optimum do not hold. Then discrete representation of alternatives needs a different approach where, instead of deriving demand functions as in consumer theory, choice theory deals directly with the utility functions.

The following section relies on Discrete Choice Analysis by Ben-Akiva and Lerman (1985), Qualitative Choice Analysis by Train (1986), Structural Analysis of Discrete Data with Econometric Applications by Manski and McFadden (1981) and Logit and Probit by Borooah (2002).

Consider an individual (decision maker) that is faced with a choice situation between a set of alternatives (i.e. choosing between organic and conventional fresh produce). Let the individual be represented by n and the choice set by C_n, the choice set is indexed by n to represent the possibility that different individuals might face different sets of alternatives in similar choice situations. Assume also that individuals have consistent and transitive preferences over the alternatives that determine a unique preference ranking. Then, an individual would obtain some relative utility from each alternative if he chooses it. Designate the utility from alternative i, an element of the choice set C_n, as U_{in}.

(8) $U_{in}, i \in C_n$

Such that alternative $i \in C_n$ is chosen if and only if:

(9) $U_{in} > U_{jn}$, all $j \neq i$, $j \in C_n$

Using the concept of indirect utility[3], the utility function can be define in terms of a vector of the attribute values for alternative i (z_{in}) and a vector of socioeconomic characteristics for individual n (S_n) that explains the variability of tastes across individuals in a specific sample. Then the utility function can be written in a general form as:

(10) $U_{in} = U(z_{in}, S_n)$

Substituting (9) in (10) provides:

(11) $U(z_{in}, S_n) > U(z_{jn}, S_n)$, all $j \neq i$, $j \in C_n$

Given this framework, the individual is assumed to select the alternative with the highest utility. However, the utilities U_{in} are not known to the researcher (analyst) with certainty. The researcher does not observe all the relevant factors and does not know the utility functions exactly. For this reason the utilities are treated as random variables. Then, the probability that the individual chooses alternative i is equal to the probability that the utility of alternative i, U_{in}, is greater than or equal to the utilities of all other alternatives in the choice set. That is,

(12) $P(i|C_n) = \Pr[U_{in} > U_{jn}, all\ j \in C_n]$

Ben-Akiva and Lerman (1985) identify four distinct sources of randomness in the utilities that justifies this distributional assumption:

[3] The indirect utility function used by the random utility theory can be understood as a more general version of the one derived in Section 3.1, where z_{in} contains other good's attributes beside prices and S_n contains other characteristics of the consumer besides income.

1) Unobserved attributes: the vector of attributes affecting the decision is incomplete; therefore the true utility function $U_{in} = U(z_{in}, S_n, z_{in}^U)$ includes an element z_{in}^U which is not observed and random, making the utility function also random.

2) Unobserved taste variations: the true utility function $U_{in} = U(z_{in}, S_n, S_n^U)$ may contain an unobserved argument S_n^U which is specific for each individual, since the variation in this term is unknown, the utility functions U_{in} is also a random variable.

3) Measurement errors and imperfect information: the true utility function might be $U_{in} = U(\tilde{z}_{in}, S_n)$ but the attributes $\tilde{z}_{in}$ are not observable. Instead z_{in} is observed which is an imperfect measurement of $\tilde{z}_{in}$ (i.e. $\tilde{z}_{in} = Z_{in} + \tilde{\varepsilon}_{in}$ where $\tilde{\varepsilon}_{in}$ is the measurement error).

4) The use of instrumental variables: the true utility function is $U_{in} = U(\hat{z}_{in}, S_n)$ but some elements of $\hat{z}_{in}$ are not observable. $\hat{z}_{in}$ is replaced with z_{in} which includes instrumental variables (i.e. $\hat{z}_{in} = g(z_{in}) + \hat{\varepsilon}_{in}$ where g denotes the imperfect relationship between the true attributes and the instrument variables and $\hat{\varepsilon}_{in}$ is a random error).

Given the presence of these sources of randomness in the utility functions, this can be considered the random utility of an alternative as a sum of observable and unobservable components of the total utilities:

(13) $U_{in} = V(z_{in}, S_n) + \varepsilon(z_{in}, S_n) = V_{in} + \varepsilon_{in}$

Where:

$V(z_{in}, S_n) = V_{in}$ = observed or systematic component of utility

$\varepsilon(z_{in}, S_n) = \varepsilon_{in}$ = unobserved or random component of utility

Replacing (13) in (12) obtains:

(14) $P(i|C_n) = \Pr[V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn}, \forall j \in C_n, j \neq i]$

According to Train (1986), the probability that person n chooses alternative i ($P(i|C_n)$), is the limit of the proportion of times, as the number of times increases without bound, that the researcher would observe a decision maker who faces the same alternative as person n, and with the same values of observed utility, to choose alternative i.

Rearranging,

(15) $P(i|C_n) = \Pr[\varepsilon_{jn} - \varepsilon_{in} < V_{in} - V_{jn}, \forall j \in C_n, j \neq i]$

The right-hand side of equation (15) is a joint cumulative distribution, the probability that the random variable $\varepsilon_{jn} - \varepsilon_{in}$ is below the known value $V_{in} - V_{jn}$ for all j in $C_n, j \neq i$. By knowing the joint probability distribution of the full set of disturbances ε's, namely $\{\varepsilon_{jn}, j \in C_n\}$ the distribution of each difference $\varepsilon_{jn} - \varepsilon_{in}$ can be derived and by using equation (15) calculate the probability that the individual will choose alternative i as a function of $V_{in} - V_{jn}$ for all j in $C_n, j \neq i$.

In general, let f (ε_{1n}, ε_{1n} ,…., ε_{Jn}) denote the joint density function of the disturbance terms, then consider alternative i to be the first alternative in C_n, then the probability that individual n will choose alternative i will be:

(16) $$P_n(1) = \int_{\varepsilon_{1n}=-\infty}^{\infty} \int_{\varepsilon_{2n}=-\infty}^{V_{1n}-V_{2n}+\varepsilon_{1n}} \cdots \int_{\varepsilon_{J_n n}=-\infty}^{V_{1n}-V_{J_n n}+\varepsilon_{1n}} f(\varepsilon_{1n}, \varepsilon_{2n}, \ldots\ldots, \varepsilon_{J_n n}) d\varepsilon_{J_n n} d\varepsilon_{J_n -1, n} \ldots\ldots d\varepsilon_{1n}$$

Different distributions of the disturbances give rise to different functional forms for the choice probabilities. The choice of the distributions (*f*) is an empirical question faced by the researcher and is addressed in the next chapter.

3.3 Empirical Specification and General Empirical Model

Following Random Utility Theory, equation (16) allows specification of an empirical model for the choice of a specific alternative (i.e. organic versus conventional produce). Consider a choice set (C_n) defined by two alternatives (J=2). The first alternative represents the choice of one product (i.e. organic produce) while the second alternative represents the choice of a competing product (i.e. conventional produce).

The individual n compares the utility (U_{ni}, i=1,2) from both alternatives and selects the alternative for which the utility is greater.

(17) $$U_{ni} = V_{in} + \varepsilon_{in} \begin{cases} U_{n1} = V_{1n} + \varepsilon_{1n} & \Rightarrow \text{Utility obtained from alternative 1} \\ U_{n2} = V_{2n} + \varepsilon_{2n} & \Rightarrow \text{Utility obtained from alternative 2} \end{cases}$$

In order to make the Random Utility Theory operational the deterministic component of the utility, V_{in}, and the distribution of the random component ε_{in} (disturbances) must be specified. Suppose that the deterministic component of the utility $V(z_{in}, S_n) = V_{in}$ is a linear function of the vector of attributes (z_{in}) and the vector of socio economic characteristics of the consumer (S_n) with this form:

$$(18)\ V_{in} = \sum_{r=1}^{R} \gamma_{nr} z_{ir} + \sum_{s=1}^{S} \beta_{is} S_{ns}$$

Where R is the number of attributes for alternative *i*, S is the number of socio-economic characteristics for consumer *n*, and γ and β are parameters to be estimated. In the area of consumer research there are many studies that have assumed that the systemic component of utility is function of only the vector of socio economic characteristics, yielding a practical and consistent with theory model called "Multinomial Logit"[4]. Among the studies using this framework in studying consumer preferences for environmental friendly consumer goods are: Byrne et. al., 1991; Byrne, Bacon and Toensmeyer, 1994; Holland and Wessells, 1998; Wessells, Johnston and Donath, 1999; Govindasamy and Italia, 1999; Teisl, Roe and Levy, 1999; Govindasamy, Italia and Adelaja, 2001; Loureiro, McCluskey and Mittelhammer, 2001.

Hence, the systematic component of utility assumed by the multinomial logit is:

$$(19)\ V_{in} = \sum_{s=1}^{S} \beta_{is} S_{ns}$$

[4] The Multinomial Logit Model is discussed in detail in the next chapter.

The multinomial logit model defines the systematic component of the utility as a function of only sociodemographic characteristics as presented by equation (19). Changing the order of the subscripts to the standard format in literature and letting the vector x_n contain the sociodemographic characteristics defined in $S_{ns,}$ equation (19') is obtained. Let K be the number of independent variables contain in x with index k.

$$(19')\quad V_{ni} = \sum_{k=1}^{K} S_{nk}\beta_{ik} = x_n B_i$$

Replacing (19') in (13):

$$(20)\quad U_{ni} = x_n \mathrm{B}_i + \varepsilon_{ni}$$

Then, the probability that the utility of organic produce is higher than the utility of conventional produce can be derived using equation (15) and (20).

$$(21)\quad \Pr(Y_n = 1) = \Pr\lfloor U_{n1} \geq U_{n2} \rfloor$$

$$= \Pr\lfloor x_n \mathrm{B}_1 + \varepsilon_{n1} \geq x_n \mathrm{B}_2 + \varepsilon_{n2} \rfloor$$

$$= \Pr[(x_n \mathrm{B}_2 + \varepsilon_{n2}) - (x_n \mathrm{B}_1 + \varepsilon_{n1}) \leq 0]$$

According to Greene (2000) an observed discrete choice variable that reveals which one provides the greater utility, but not the observable utilities can be defined. A discrete dependent variable representing the observed outcome of a binary choice is defined as:

$$(22)\; y_{ni} = \begin{cases} 1 & \text{if individual } n \text{ prefer alternative 1 (i=1)},\; U_{n1} \geq U_{n2} \\ 0 & \text{if individual } n \text{ prefer alternative 2 (i=2)},\; U_{n1} \leq U_{n2} \end{cases}$$

The empirical specification of the utility levels underlying the discrete choice model has the following form:

(23) $\Pr(y_{ni} = 1) = F(x_n \mathrm{B}_i + \varepsilon_{ni})$

Where x is a vector of socio-demographic characteristics, B is a vector of parameters to be estimated, F is a cumulative distribution function and ε is an error term. The methodology used to calculate the parameters depend on the assumptions made on the distribution of the disturbances (F). As explained at the end of the last section, this is an empirical question that will be addressed in the next chapter.

4. Approach and Methodology

The principal objective of this study is to determine the factors that influence consumers to choose organic over conventional fresh produce. Discrete choice models require information about consumer's preferences and socio-demographic characteristics, thus a survey was used as the principal tool for data collection. The survey design, the trial runs, and the data collection process are described in the next sections.

Information regarding attitudes and behavior towards the environment and health were collected by using a large number of behavioral questions that were included in the survey. To condense the information and make it useable for econometric estimation a factor analysis was used. Factor analysis and the behavioral questions used in this study are described in Section 4.3. The final Section of this Chapter covers the multinomial logit technique which is the method chosen to estimate the discrete choice model.

4.1 Survey Design and Trial Runs

The survey used in this study was designed to be user-friendly and flexible enough to collect the data needed. With the availability of new technologies to collect and store information, this study made use of information technologies to design a survey that can be placed on and collect data through a computer interface. In this context, the survey design was in part result of the specific characteristics that these technologies offered to the researchers.

A test survey was designed using the HTML language (Hypertext Markup Language) and placed on a server that allowed respondents to access the survey through the Internet. A sample of 160 students was chosen from undergraduate classes in the Agricultural and Resource Economics Department (AREC). A total of 117 (73%) valid responses were obtained. The survey allowed the researchers to test the questions aimed to explore the attitudes of the students towards the environment and their health, their fresh produce preferences and basic demographic information. The information from this survey also allowed testing of the behavioral questions about the environment and health, and pre-testing of the factor analysis and the econometric model proposed in the next chapter. From the results obtained by the test run over the internet, the survey design was improved, new questions were added and others were excluded in order to maintain a simple and friendly design.

The final survey used in the data collection in this research is presented in Appendix 1. The final version of the survey had 31 questions and was expected to be completed in no more than 20 minutes. Subsequent analysis showed that the average time needed to complete the survey was around 10 minutes. The survey covers four areas: one to collect information about their fresh produce shopping habits and their preference between organic and conventional-grown for a set of products that included fresh produce, drinks, meat and poultry, and others. The next section was designed to collect information about the willingness to pay a premium for organic fruits and vegetables. The third part contained the 29 behavioral questions related to the environment and health behavior of the respondent. The final part of the survey contained demographic questions including age, income, education, gender and others. All the questions were close ended; multiple choice and scale type questions were used. None of the questions were open to simplify the data collection process as explained below.

4.2 Data Collection

The target audience (population) for this study was shoppers aged 18 and over who live in the Portland, Oregon metropolitan area. A supermarket, a food co-operative and a farmer's market were chosen to collect responses on consumer's choice of organic produce. These institutions were approached for approval for interviewing shoppers at their facilities. The respondents were randomly selected and were offered a compensation of a five dollar ($5) certificate to be used in the store or market. There were no limitations on recruitment based on gender, race or ethnicity.

This specific population was selected because they face purchasing decisions regarding produce at these establishments. This sample might not reflect the characteristics of the population and may also have some degree of choice based bias. The sampling is deliberately skewed in favor of observing organic consumers in order to achieve a more balanced sample with respect to organic purchasing. Caution should be taken in transmitting the results obtained from this sample to the general population.

Data collection took place through the HTML survey that was loaded on five Tablet PCs equipped with touch screens that were taken to the chosen locations. The Tablet PCs are small computers that make their transportation and use practical and cost efficient. Moreover, the Tablet PCs and their touch screens allow the respondents to quickly answer the survey questions. Placement of the computers varied by venue. At the farmer's market the booth was placed on a corner using two tables at a right angle to each other, with chairs to sit on facing inwards, and passers by in each direction. At the conventional supermarket tables were placed in a row outside the front door of the store to avoid interrupting the normal activities in the location. At the co-operative the booth was placed inside the store after the checkout. In all cases a laminated poster advertising the survey was placed with the text "Fill out a survey about produce choices, And earn a $5 Certificate." Other small posters and flyers were also posted in different parts of the stores (i.e. the produce section, checkout line). The posters also contained the university logo and the name of the experiment station. A team of 2-3 individuals randomly recruited respondents from the shoppers mentioning the compensation payment of 5 dollars for shopping at that market. The farmer's market was the easiest location at which to recruit, eventually people lined up to take the survey. The conventional supermarket was the hardest location at which to recruit, though because of higher customer traffic, it did not take as long as at the cooperative to achieve the 100 survey goal.

The recruitment process included the following steps: 1) Shoppers entering the store were asked randomly to participate in the survey by reading an introductory script. 2) If the shopper agreed to participate in the study, he/she was directed to an open Tablet PC. An assistant was at the table at all times to provide support and answer any questions asked by the respondents. 3) The survey started with an informed consent page. If consent was indicated, the program took them to the main survey. Neither the name of the shopper or any contact information (e.g. address or telephone number) that could lead to the identification of the respondent was requested. 4) After the completion of the survey the respondents received the compensation payment.

As mentioned earlier, three market outlet types were chosen to collect the data. Since the principal purpose of this research is to study factors affecting the choice of organic and conventional produce, it was critical to choose shopping locations that offer these two types of fresh produce. The three locations, the supermarket, the farmer's market and the co-operative offered a wide selection of fruits and vegetables that included organic, conventional and eco-labeled products.

The supermarket where the survey was taken is located in a suburb of the Portland metro area. The store belongs to a regional supermarket chain with local ownership and many locations in the metropolitan area. The farmer's market is located in downtown Portland and offers a wide variety of locally grown products and draws in regular as well as occasional shoppers. Finally, the food co-operative is a small/medium sized grocery store located in Portland and offers a wide variety of organic and natural products. The data collection was conducted on a Saturday or Sunday at each location, over a 4 week period. In total, 300 surveys were completed-100 at each location.

4.3 Factor Analysis

As explained in section 4.1, the survey included 29 questions related to environmental and health behavior and attitudes of the respondent. These questions contained information about consumer's motives for demanding organic produce. Factor analysis was used to condense the survey data gathered from these questions into a form useable for econometric estimation of the factors influencing choice of organic produce. It is not practical to use 29 variables in the econometric estimation of a discrete choice model to approach only two issues (environment and health). Factor analysis offers a method to condense information from a set of many variables into a few variables called "factors". This section begins with a brief discussion about the behavioral questions used in this study followed by the methodological details of the factor analysis model.

4.3.1 Behavioral Questions

One of the goals of this research is to determine the relative importance and magnitude of the effects of environmental and health motives in the demand for organic produce. The current literature about empirical research of consumer preferences towards environmental friendly products have made use of "attitudinal questions" in eliciting consumer's concern of environmental and health related issues. Another way to approach the same problem is the use of "behavioral questions" in collecting information about the behavior and perceptions of the consumer towards these issues.

Behavioral questions have been widely used in other social sciences and business related fields. Particularly, marketing research has used behavioral questions intensively in order to segment markets. Behavioral questions seek factual information with respect to what people do, how often, with whom; what people own, how they use products, etc; while attitudinal questions seek information about people's beliefs, opinions, motivations, images, intentions, etc, with respect to products, services, ideas, etc.

Roberts (1996) reports that there is a gap between attitude and behavior expressed by consumers in the market of environmentally friendly products. He reviews the results conducted by Simmons Market Research Bureau in which it was found that people in the U.S. do not actually buy the products they claim to prefer. He adds that high concern over the environment was found, but behaviors consistent with such concern were lacking. Quoting Roberts:

> *"Ultimately we must investigate consumer behavior because it is such behavior, not expressed concern, that will help correct the problems currently facing the environment and create markets for green products and services."*

Among the reasons he attributes the attitude-behavior gap to in green markets are: 1) Green products are much more expensive than conventional products; 2) Price, quality and convenience still being the most important attributes, having the green appeal might be a competitive advantage if the first ones are not compromised; and 3) Consumers are confused about green products.

This research uses a set of behavioral questions found in the marketing research field to elicit consumer's behavior toward the environment and their health in order to identify factors that will explain the choice of organic produce. Moreover, this study will use the methodology of "Factor Analysis" to process, analyze and reduce the information collected by the behavioral questions.

The environmental behavioral questions used in this analysis are part of a much bigger set that belongs to what is called "Environmental Conscious Consumer Behavior" (ECCB) in the marketing literature. Specifically, the questions were taken from the studies conducted by Roberts (1996) and Straughan and Roberts (1999) which are the most comprehensive studies on ECCB to date. ECCB measures the extent to which an individual purchases goods and services believed to have a more positive impact on the environment. ECCB makes use of behavioral questions in which the response categories are in 5-point Likert-type format that goes from "always true" to "never true". According to Roberts (1996), ecologically conscious consumers are defined as "those who purchased products and services which they perceive to have a positive impact on the environment". The complete list of the ECCB questions used in the survey is presented in Appendix 4.

In the case of health related questions, they were taken from a study by Kraft and Goodell (1993) in which they create a "wellness scale" to identify the health conscious consumer. The wellness concept used by these authors consisted in a holistic approach for improving the quality of health and life. The scale measures individual's interests in their health and in their behavior aimed at maintaining or improving it. The scale identify five dimensions of wellness: self-responsibility, nutritional awareness, stress awareness and management, physical fitness and environmental sensitivity. They reported 19 questions that can be used to create the wellness scale and identify health-related segments. These questions include some about beliefs and interests as well as behavior. The questions used by this research are presented in Appendix 5.

Thirteen environmental and fifteen health related plus one additional variable regarding the importance of free of Genetically Modified Organisms (GMOs) for consumers[5] are used. The topics covered by the questions are very broad. It is necessary to identify the most important ones in order to create a reduced set of vectors that explains environmental and health consumer's behavior. One methodology suited for the task is factor analysis.

4.3.2 Factor analysis Model - The Principal Components Approach

Factor Analysis consists of a number of statistical techniques aimed at simplifying complex sets of data. Specifically, factor analysis is used in the social sciences to simplify correlation matrices by identifying common "factors" related with a set of variables. Factor analysis has been used in a wide number of fields, including Psychology, Chemistry, and Business. The two goals of this technique are to discover or to reduce the dimensionality of the data set and to identify new meaningful underlying variables. Factor analysis attempts to identify underlying variables that explain the pattern of correlations within a set of observed variables. Factor analysis

[5] This question asked consumers how important is obtaining free of GMO products when buying fresh produce. This variable was included since GMO has become an important issue when taking about organic food.

is often used to identify a small number of factors that explain most of the variance observed in a much larger number of variables.

According to Kline (1994) there are many methods to approach factor analysis, two of the most important are Principal Components and Principal Axes Methods. The difference between the two lies on specific assumptions in the construction of the correlation matrix. In this study principal components has been chosen for the simplicity of the construction of the correlation matrix and the close resemblance between the two methods in large sets of data. Kline reports that in a 1976 study by Hartman, it was found that with large sets of data principal components and principal axes methods were identical.

The next section relies on "Methods of Multivariate Analysis" by Srivastava (2002), "Methods of Multivariate Analysis" by Rencher (2002), and "Handbook of Statistics 14: Statistical Methods in Finance" by Maddala and Rao (1996).

Consider a multivariate data matrix $y = (y_1, y_2, \ldots, y_p)'$ with n rows (sample size) and p columns (number of variables). The p elements of each row are scores or measurements on the variables of interest (i.e. ratings of the 13 environmental and 15 health related behavioral questions). In factor analysis the goal is to represent the variables $y_1, y_2, \ldots, y_p$, as linear combinations of a few random variables $f_1, f_2, \ldots, f_m$, $(m < p)$ called *factors*. According to Kline (1994), a factor is a dimension or construct which is a condensed statement of the relationships between variables. Like the original variables, the factors vary from individual to individual but they cannot be observed since they are hypothetical variables.

If the original variables in $y = (y_1,y_2,...,y_p)'$ are correlated, the dimension of the system is less than p. The goal of factor analysis is to reduce the redundancy among the variables by using a smaller number of factors (i.e. reducing the rank of the matrix y). If there are subsets of variables in y that are highly correlated among the variables in a specific subset but with low correlations between a subset and another, then may be a few underlying factors gave rise to the variables in the subset.

Factor analysis expresses each variable $(y_1,y_2,...,y_p)$ as a linear combination of underlying factors $(f_1, f_2,..., f_m)$ including an error term (ε) to account for that part of the variable that is unique and not in common with the other variables. That is:

$$
\begin{aligned}
&y_1 - \mu_1 = \lambda_{11} f_1 + \lambda_{12} f_2 + ... + \lambda_{1m} f_m + \varepsilon_1 \\
(24)\quad &y_2 - \mu_2 = \lambda_{21} f_1 + \lambda_{22} f_2 + ... + \lambda_{2m} f_m + \varepsilon_2 \\
&\vdots \\
&y_p - \mu_p = \lambda_{p1} f_1 + \lambda_{p2} f_2 + ... + \lambda_{pm} f_m + \varepsilon_p
\end{aligned}
$$

The coefficients λ_{ij} are called loadings and are weights on how each variable y_i depends on the factors. In other words they show the importance of the j^{th} factor f_j to the i^{th} variable y_i. After estimating the factor loadings and rotating them (explained later), the factor analysis is intended to partition the variables into groups corresponding to factors. Then instead of using p explanatory variables m factors ($m < p$) can be used.

Using matrix notation, the system of equation (24) can be represented as:

(25) $y - \mu = \Lambda f + \varepsilon$

Where $y = (y_1, y_2,..., y_p)'$, $\mu = (\mu_1, \mu_2,..., \mu_p)'$, $f = (f_1, f_2,..., f_m)'$, $\varepsilon = (\varepsilon_1, \varepsilon_2,..., \varepsilon_p)'$,

and :

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{bmatrix}$$

Assumptions of Factor Analysis:

1) Assume that $E(f_j) = 0$, in matrix notation $E(f)=0$.

2) $Var(f_j) = 1$, and $Cov(f_j, f_k)=0$ for $j \neq k$.. Equal to $Cov(f)=I$.

3) $E(\varepsilon_i) = 0$, $E(\varepsilon) = 0$ in matrix notation.

4) $Var(\varepsilon_i) = \psi_i$, and $Cov(\varepsilon_i, \varepsilon_k)=0$ for $i \neq k$. Equal to $Cov(\varepsilon)=\Psi$.

5) $Cov(\varepsilon_i, f_j)=0$ for all i and j. $Cov(f, \varepsilon)= O$.

With these assumptions the variance of each y_i is equal to:

(26) $Var(y_i) = \sigma_{ii} = \lambda_{i1}^2 + \lambda_{i2}^2 + \ldots + \lambda_{i\,m}^2 + \psi_i$

Or in matrix terms:

(27) $Cov(y) = \Sigma = Cov(\Lambda f + \varepsilon)$

$$\begin{aligned} &= \Sigma = Cov(\Lambda f) + Cov(\varepsilon) \\ &= \Sigma = \Lambda Cov(f) \Lambda' + \Psi \\ &= \Sigma = \Lambda I \Lambda' + \Psi \\ &= \Sigma = \Lambda \Lambda' + \Psi \end{aligned}$$

Equation (27) represents a simplified structure of the Covariance matrix for the original variables y, since this structure depends only on the factor loadings for m factors. It is extremely important to note that the covariance of the variables with the factors is equal to the factor loadings themselves. The covariance between y_1 and f_2, for example, is equal to:

$$Cov(y_1, f_2) = E[(y_1 - \mu_1)(f_2 - \mu f_2)]$$

$$\begin{aligned}
&= E[(\lambda_{11} f_1 + \lambda_{12} f_2 + \ldots + \lambda_{1m} f_m)(f_2)] \\
&= E(\lambda_{11} f_1 f_2 + \lambda_{12} f_2^2 + \ldots + \lambda_{1m} f_m f_2) \\
&= \lambda_{11} Cov(f_{1,} f_2) + \lambda_{12} Var(f_2) + \ldots + \lambda_{1m} Cov(f_m, f_2) \\
&= \lambda_{12}
\end{aligned}$$

In general terms:

$$(28)\quad \begin{matrix} Cov(y_i, f_j) = \lambda_{ij} \\ Cov(y, f) = \Lambda \end{matrix} \quad i=1,2,\ldots,p\ ; j=1,2,\ldots,m$$

Equation (28) can be divided into two components:

$$(29)\quad \begin{aligned} Var(y_i) = \sigma_{ii} &= (\lambda_{i1}^2 + \lambda_{i2}^2 + \ldots + \lambda_{i\,m}^2) + \psi_i \\ &= h_i^2 + \psi_i \\ &= communality(h_i^2) + specific\,\mathrm{var}iance(\psi_i) \end{aligned}$$

The communality, also called the common variance, is the proportion of the variance which can be explained by common factors. On the other hand, specific variance is particular to each variable.

The next step in Factor Analysis is to estimate the factor loadings and the communalities of Equation (29). One of the techniques available is called *principal components method*. Given a random sample $(y_1, y_2, \ldots, y_n)$, the sample covariance matrix **S** is estimated and an estimator for Λ, denoted by $\hat{\Lambda}$, is found. Then the sample approximation of Equation (29) is:

$$(30)\quad S \cong \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$$

Consider only the communalities in the sample estimation of the covariance matrix for the moment. ($S \cong \hat{\Lambda}\hat{\Lambda}'$). In order to factor S spectral decomposition of S is undertaken:

(31) $S = CDC'$

Where C is an orthogonal matrix constructed with normalized eigenvectors $(c'_i c_i)=1$ of S as columns and D is a diagonal matrix with the eigenvalues of S $(\theta_1, \theta_2, \ldots, \theta_p)$. Since the matrix S is variance-covariance matrix, it is positive semidefinite with eigenvalues all positive or zero, D can be factored as $D=D^{1/2}D^{1/2}$. Then Matrix S becomes:

(32) $S = CDC' = CD^{1/2}D^{1/2}C'$

$= (CD^{1/2})(CD^{1/2})'$

Define $D_1=\text{diag}(\theta_1, \theta_2, \ldots, \theta_m)$ with the *m* largest eigenvalues $\theta_1 > \theta_2 > \ldots > \theta_p$ and $C_1 = (c_1, c_2, \ldots, c_m)$ containing the corresponding eigenvectors. Λ can be estimated from the first *m* columns of $CD^{1/2}$, as:

(33) $\hat{\Lambda} = C_1 D_1^{1/2} = (\sqrt{\theta_1 c_1}, \sqrt{\theta_2 c_2}, \ldots, \sqrt{\theta_m c_m})$

Where $\hat{\Lambda}$ is a *p* x *m* matrix, C_1 is *p* x *m* and $D_1^{1/2}$ is *m* x *m*. To complete the approximation of S an estimator for ψ is needed. This estimator is given by:

(34) $\hat{\psi}_i = s_{ii} - \sum_{j=1}^{m} \hat{\lambda}_{ij}^2$

(35) $\hat{\Psi} = \text{diag}(\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_p)$

Given equations (27), (33) and (35) the matrix of factor loadings and communalities can be estimated. However there is one question still to be answered. How many factors (m) should be chosen?. Three criteria are presented by Rencher (2002):

1) Choose m equal to the number of factors necessary for the variance accounted for to achieve a predetermined percentage.
2) Choose m equal to the number of eigenvalues greater than the average eigenvalue.
3) Use a scree test based on a plot of eigenvalues of S. If the graph drops sharply, followed by a line with much smaller slope, choose m equal to the number of eigenvalues before the straight line begins.

Typically, only the first m principal components with eigenvalues greater than 1 are included in the analysis.

Following Kline (1994) there are some characteristics of the components and factor loadings that are worth noting:

1) The square summed of loadings in the rows of a factor matrix are the communalities. They indicate the proportion of variance in each variable which the factors explain. This is given by: $\hat{h}_i^2 = \sum_{j=1}^{m} \hat{\lambda}_{ij}^2$.
2) The sum of the squared loadings of each factor (column) is equal to the eigenvalue of that factor. Which is the variance explained by each factor.

$\sum_{i=1}^{p} \hat{\lambda}_{ij}^2 = \theta_j$. And the average of the squared loadings of a factor shows the percentage of variance in the correlation matrix explained by that factor.

$$\frac{\sum_{i=1}^{p} \hat{\lambda}_{ij}^2}{tr(S)} = \frac{\theta_j}{tr(S)}.$$

3) When factors are uncorrelated, as is the case of principal component analysis, factor loadings are not only the correlation but the weights for predicting the variable from the factor.
4) Principal components, as a result of its computational algebra, produces an arbitrary general factor with high loadings followed by bipolar factors which makes interpretation difficult.

Because of this last property, methods of simplifying principal component analyses have been developed. One of these methods is the rotation of factors which can be carried out by many techniques. The rotation of the factor loadings is equivalent to the multiplication of $\hat{\Lambda}$ by an orthogonal matrix which preserves the essential properties of the original loadings reproducing the covariance matrix and satisfying all the assumptions. Let T be an orthogonal matrix, then $\hat{\Lambda}^* = \hat{\Lambda}T$, since TT'=I, the rotated loadings provide the same estimates of the covariance matrix as before:

(36) $S \cong \hat{\Lambda}^* \hat{\Lambda}^{*\prime} + \hat{\Psi} = \hat{\Lambda}TT'\hat{\Lambda}' + \hat{\Psi} = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$

One of the most popular rotation techniques is the Varimax rotation of factors in which the sum of variances of squared loadings in the columns of the factor matrix

($\hat{\Lambda}^*$) is maximized. This produces in each column, loadings that are either high or near zero, making the interpretation of the principal components easier. Hence the factors loadings are calculated one more time after the rotation of the factors have been done making the process of identifying and interpreting the factors easier.

The final step in Factor Analysis is to estimate the factor scores defined as $\hat{F}_i = (\hat{f}_{i1}, \hat{f}_{i2}, \ldots, \hat{f}_{im},)'$, i=1,2,…,n. Which are estimates of the underlying factor values for each observation. Since the *f*'s are not observed, it is necessary to estimate them as functions of the observed variables (y). One way of estimating the factor scores is using a centered regression model given by:

$$
\begin{aligned}
& f_1 = \beta_{11}(y_1 - \bar{y}_1) + \beta_{12}(y_2 - \bar{y}_2) + \ldots + \beta_{1p}(y_p - \bar{y}_p) + \phi_1 \\
(37)\quad & f_2 = \beta_{21}(y_1 - \bar{y}_1) + \beta_{22}(y_2 - \bar{y}_2) + \ldots + \beta_{2p}(y_p - \bar{y}_p) + \phi_2 \\
& \vdots \\
& f_m = \beta_{m1}(y_1 - \bar{y}_1) + \beta_{m2}(y_2 - \bar{y}_2) + \ldots + \beta_{mp}(y_p - \bar{y}_p) + \phi_m
\end{aligned}
$$

This is equivalent to the next system of equations in matrix form:

(37') $F = Y_c \mathrm{B}_1 + \Phi$

Solving this system of multiple equations, the estimate for B_1 is[6]:

(38) $\hat{\mathrm{B}}_1 = (Y_c'Y_c)^{-1} Y_c'F = S^{-1}\hat{\Lambda}$

Then the predicted values for f_i, the factor scores, are given by:

(39) $\hat{F} = Y_c \hat{B}_1 = Y_c S^{-1} \hat{\Lambda}$

In this study, the principal components method of data reduction is applied to the set of environmental and health behavioral questions independently, in order to identify the principal factors in each subset and calculate the factor scores for each

[6] Remember from the discussion of Principal Components that $Var(y) = S$ and $Cov(y,f) = \Lambda$.

factor. Hence, $\hat{F}^{E}{}_{n} = (\hat{f}^{E}{}_{n1}, \hat{f}^{E}{}_{n2},..., \hat{f}^{E}{}_{nm},)'$, n=1,2,…,N (N=sample size) denotes the factor scores obtained from applying factor analysis to the set of environmental questions from m factors and $\hat{F}^{H}{}_{n} = (\hat{f}^{H}{}_{n1}, \hat{f}^{H}{}_{n2},..., \hat{f}^{H}{}_{no},)'$, n=1,2,…,N denotes the scores obtained from the set of health questions obtained from o factors. The principal component analysis is carried out using SPSS for windows using a Varimax rotation of factors with eigenvalues greater than 1 that have passed the scree test.

4.4 Multinomial Logit Model

The discrete choice model of section 3.3 needs a specification of the disturbances in order to be empirically estimated. The multinomial logit offers a way of deriving an econometric model based on the grounds of the random utility theory discussed in section 3.2. This section presents the derivation of the multinomial logit for discrete choice models, the estimation methods required and the statistical tests used in these models.

Recalling from section 3.2, the Random Utility framework yields equation (13) which divided the random variables U_{in} into a systematic part of the utility function (V_{in}) and a random component (ε_{in}). From (13) derive Equation (15) which states that the probability that alternative i is chosen is equal to probability that the difference between the disturbances is less or equal than the difference between the systematic component of the utility between alternative i and the other alternatives.

(13) $$U_{in} = V(z_{in}, S_{n}) + \varepsilon(z_{in}, S_{n}) = V_{in} + \varepsilon_{in}$$

(15) $P(i|C_n) = \Pr[\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n, j \neq i]$

If alternative i is chosen then $U_{in} \geq U_{jn}$, for all $j \neq i$, $j \in C_n$, which is equivalent to $U_{in} \geq Max U_{jn}$, for all $j \neq i$, $j \in C_n$. The right hand side of this equation is in fact a "composite" alternative out of all the elements in C_n other than i. Then equation (14) and (15) can be represented as:

(40) $$P(i|C_n) = \Pr\left[V_{in} + \varepsilon_{in} \geq \max_{j \in C_n j \neq i}(V_{jn} + \varepsilon_{jn})\right]$$

Where the composite alternative of the right hand side is a random variable, since all U_{jn} for $j \neq i$ are random variables. In order to derive the multinomial logit it is necessary to derive the distribution of the composite alternative from the underlying distribution of the disturbances.

The Multinomial Logit assumes that the disturbances ε_{in} are independently and identically distributed following a Gumbel distribution with a location parameter η, and a scale parameter μ>0. If ε_{in} is Gumbel-distributed. Then:

(41) CDF : $F(\varepsilon) = \exp[-e^{-\mu(\varepsilon - \eta)}], \mu > 0$

(42) PDF: $f(\varepsilon) = \mu e^{-\mu(\varepsilon - \eta)} \exp[-e^{-\mu(\varepsilon - \eta)}]$

Following Ben-Akiva (1985) the Gumbel distribution has the following properties:

1) The mode is η.

2) The mean is $\eta + \gamma/\mu$, where γ is Euler constant ~ 0.577.

3) The variance is $\pi^2/6\mu^2$.

4) If ε_{in} is Gumbel distributed with parameter (η,μ) while V, and α>0 are any scalar constants, then αε + V is Gumbel distributed with parameters (αη + V, μ/α).

5) If ε_1 and ε_2 are independent Gumbel-distributed variates with parameters (η_1,μ) and (η_2,μ), respectively, then ε* = ε_1 - ε_2 is logistically distributed as:

$$(43)\ F(\varepsilon) = \frac{1}{1 + e^{\mu(\eta_2 - \eta_1 - \varepsilon^*)}}$$

6) If ε_1 and ε_2 are independent Gumbel-distributed with parameters (η_1,μ) and (η_2,μ), respectively, then max (ε_1,ε_2) is Gumbel-distributed with parameters $\left(\frac{1}{\mu}\ln(e^{\mu\eta_1} + e^{\mu\eta_2}), \mu\right)$.

7) If (ε_1, ε_2,..., ε_J) are J independent Gumbel-distributed random variables with parameters (η_1,μ) , (η_2,μ),..., (η_J,μ) respectively, then max (ε_1, ε_2,..., ε_J) is Gumbel distributed with parameters $\left(\frac{1}{\mu}\ln\sum_{j=1}^{J} e^{\mu\eta_j}, \mu\right)$.

Without loss of generality, one can assume alternative *i* to be the first alternative in C_n, then equation (40) is :

$$(44)\ P_n(1) = \Pr\left[V_{1n} + \varepsilon_{1n} \geq \max_{j=2,\ldots,J_n}(V_{jn} + \varepsilon_{jn})\right]$$

Define

$$U_n^* = \max_{j=2,\ldots,J_n}(V_{jn} + \varepsilon_{jn})$$

By using Property 7, U_n^* is Gumbel-distributed with parameters $\left(\frac{1}{\mu}\ln\sum_{j=2}^{Jn} e^{\mu V_{jn}}, \mu\right)$.

From Property 4, obtain $U_n^* = V_n^* + \varepsilon_n^*$, where

$$(45)\ V_n^* = \frac{1}{\mu} \ln \sum_{j=2}^{Jn} e^{\mu V_{jn}}$$

And ε_n^* is Gumbel distributed with parameters $(0,\mu)$.

Replacing (24) in (20) obtain:

$$(46)\quad P_n(1) = \Pr\left[V_{1n} + \varepsilon_{1n} \geq V_n^* + \varepsilon_n^*\right]$$

$$= \Pr\left[(V_n^* + \varepsilon_n^*) - (V_{1n} + \varepsilon_{1n}) \leq 0\right]$$

by Property 5:

$$P_n(1) = \frac{1}{1 + e^{\mu(V_n^* - V_{1n})}}$$

$$= \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\mu V_n^*}}$$

$$= \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + \exp(\ln \sum_{j=2}^{Jn} e^{\mu V_{jn}})}$$

$$(47)\ P_n(1) = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^{Jn} e^{\mu V_{jn}}}$$

Which defines a proper probability mass function since[7] :

$0 \leq P_n(i) \leq 1$ for all $i \in C_n$ and

$$\sum_{i \in C_n} P_n(i) = 1$$

[7] The parameter μ is not identifiable; the usual procedure is to set it arbitrarily to a convenient value, such as 1.

The multinomial logit model defines the systematic component of the utility as a function of only sociodemographic characteristics as presented by equation (19). Changing the order of the subscripts to the standard format in literature and letting the vector x_n contain the sociodemographic characteristics defined in $S_{ns,}$ equation (19') is obtained. Let K be the number of independent variables contain in x with index k.

$$(19')\quad V_{ni} = \sum_{k=1}^{K} S_{nk}\beta_{ik} = x_n B_i$$

Replacing (19') in (13):

$$(48)\quad U_{ni} = x_n \mathrm{B}_i + \varepsilon_{ni}$$

Where B_i is Kx1 vector of parameters to be estimated indexed by alternative i, while x_n is a 1xK vector of sociodemographic characteristics indexed by consumer n, and K being the number of characteristics. Then the probability of individual n choosing alternative i given x_n is denoted by a variable $Y_n = i$:

$$(49)\quad \Pr(Y_n = i | x_n) = \frac{\exp(x_n \mathrm{B}_i)}{\sum_{j=1}^{Jn} \exp(x_n \mathrm{B}_j)}, \quad i, j = 1, 2, \ldots, J_n$$

Where J_n = number of alternatives indexed by n and n = the number of consumers. Model (49) is unindentified since more than one set of parameter generates the same probabilities. To identify the model a constraint must be imposed (Long 1997). One of the techniques available is to choose arbitrarily one of the B_j and set it equal to zero. In model (49) any B_j, $j=1,2,\ldots,J_n$ can be chosen to meet this restriction.

4.4.1 Estimation Methods and Inference

Given a random sample of N observations taken from a population, the goal is to find estimates for β from the discrete choice model presented by (49). The regression model has a non-linear relationship between the dependent variables and the parameters to be estimated. In such cases, discrete choice models are estimated using the Maximum Likelihood Estimation method (MLE). Since the observations are drawn at random from the population, the likelihood of the entire sample is the product of the likelihoods of the individual observations. The probability for each individual [$\Pr(Y_n=i)$] is given by equation (49). Define L as the likelihood function, obtain:

$$(50)\ L=\prod_{n=1}^{N}\prod_{i=1}^{J}[\Pr(Y_n=i)]^{y_{ni}}$$

Taking the Logarithm of equation (50) obtain the Log-Likelihood function:

$$(51)\ \log L=\sum_{n=1}^{N}\sum_{i=1}^{J}y_{ni}\Pr(Y_n=i)$$

Setting the first derivatives of *LogL* with respect to the coefficients equal to zero, the necessary first order conditions (Ben-Akiva 1985) are:

$$(52)\ \frac{\partial \log L}{\partial \beta_k}=\sum_{n=1}^{N}\sum_{i}[y_{ni}-\Pr(Y_n=i)]\,x_{ink}=0 \quad \text{for } k=1,\ldots,K$$

The likelihood function is globally concave which ensures the uniqueness of the MLE estimates (Long 1997). Since (50) is a non-linear model, numerical methods are used to find estimates that maximize the log-likelihood function according to the first order conditions presented by (52). The Newton-Raphson Method is the numerical method used to solve model (50) (Long 1997). Consider $\theta = (B_1, B_2, \ldots\ldots B_K)'$ be the vector of the parameters of the Log-Likelihood function. The Hessian Matrix is:

$$(53)\ \frac{\partial^2 \ln L}{\partial\theta\partial\theta'} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial\beta_1\partial\beta_1} & \cdots & \frac{\partial^2 \ln L}{\partial\beta_1\partial\beta_k} \\ \vdots & & \vdots \\ \frac{\partial^2 \ln L}{\partial\beta_K\partial\beta_1} & \cdots & \frac{\partial^2 \ln L}{\partial\beta_K\partial\beta_K} \end{pmatrix}$$

The Newton-Raphson algorithm proceeds according to the equation (54). Iterations continue until there is convergence, that is when the gradient of the Log-likelihood is close to 0 and the estimates do not change from one step to another:

$$(54)\ \theta_{n+1} = \theta_n - \left(\frac{\partial^2 \ln L}{\partial\theta\partial\theta'}\right)^{-1} \frac{\partial \ln L}{\partial\theta_n}$$

The Newton-Raphson Method also provides estimates for the asymptotic covariance matrix Var($\hat{\theta}$) which is equal to the inverse of the negative of the expected value of the Hessian, also known as the information matrix, given by:

$$(55)\ Var(\hat{\theta}) = \left(-E\left[\frac{\partial^2 \ln L}{\partial\theta\partial\theta'}\right]\right)^{-1} = I(\theta)^{-1}$$

The Maximum Likelihood Estimators obtained by maximizing (51) have the following properties (Greene 2002):

1) It is Consistent: plim $\hat{\theta} = \theta$
2) It is asymptotically normally distributed: $\hat{\theta} \overset{a}{\sim} N[\theta, I\{\theta\}^{-1}]$
3) It is asymptotically efficient: no other consistent asymptotically normally distributed estimator has a smaller covariance matrix.

4) MLE estimators are invariant: the maximum likelihood estimators of any continuous function of B is that function of the MLE estimator.

The marginal effect for the k^{th} continuous variable, defined as the partial change in the probability of an event, is obtained by taking the derivative of (48) with respect to k^{th} variable (Long 1997). The marginal effect is the slope of the curve relating x_k to $\Pr(y_n=i|x_n)$ holding all other variables constant :

(56) $$\frac{\partial \Pr(Y_n = i|x_n)}{\partial x_k} = \Pr(Y_n = i|x_n)\left[\beta_{ki} - \sum_{j=1}^{2} \beta_{kj} \Pr(Y_n = j|x_n)\right]$$

Another form to express the marginal effect for continuous variables when there are 2 alternatives in the choice set is:

(57) $$\Pr(y_{ni} = 1|x_n) = F(x_n \mathrm{B}_i)$$

(58) $$\frac{\partial F(x_n\mathrm{B}_i)}{\partial x_k} = \frac{\partial F(x_n\mathrm{B}_i)}{\partial x_n\mathrm{B}_i}\frac{\partial x_n\mathrm{B}_i}{\partial x_k} = f(x_n\mathrm{B}_i)\beta_k = \frac{\exp(x_n\mathrm{B}_i)}{[1+\exp(x_n\mathrm{B}_i)]^2}\beta_k$$

(59) $$= \Pr(y_{ni} = 1|x_n)[1 - \Pr(y_{ni} = 1|x_n)]\beta_k$$

Where F is the CDF and f is the PDF for the Logistic Distribution. Since the value of the marginal effect depends on the levels of all the variables contained in x_{ni} it is necessary to decide on which values of the variables to use when compute the marginals. This study reports two methods of calculating the marginal effects, at the average over all observations and at the mean of the independent variables.

(60) Average of all observations: $$\frac{\partial \Pr(y_{ni} = 1|x_n)}{\partial x_k} = \frac{1}{N}\sum_{n=1}^{N} f(x_n\mathrm{B}_i)\beta_k$$

(61) At the means: $$\frac{\partial \Pr(y_{ni} = 1|x_n)}{\partial x_k} = f(\bar{x}_n\mathrm{B}_i)\beta_k$$

The marginal effects for continuous variables are also calculated examining an unit increase that is centered around $\bar{x}_k$. This is interpreted as an unit change in x_k that is centered around x_k results in a change of in the predicted probability $\frac{\Delta \Pr(y_{ni}=1|\bar{x}_n)}{\Delta x_k}$, holding all other variables at their means:

$$(62)\quad \frac{\Delta \Pr(y_{ni}=1|\bar{x}_n)}{\Delta x_k} = \Pr(y_{ni}=1|\bar{x}_n, \bar{x}_k + 1/2) - \Pr(y_{ni}=1|\bar{x}_n, \bar{x}_k - 1/2)$$

The marginal effects for discrete variables (dummies) are calculated examining the change in probability observed from changing the discrete variable from 0 to 1 holding the rest of the variables at their means.

$$(63)\quad \frac{\Delta \Pr(y_{ni}=1|\bar{x}_n)}{\Delta x_k} = \Pr(y_{ni}=1|\bar{x}_n, x_k = 1) - \Pr(y_{ni}=1|\bar{x}_n, x_k = 0)$$

In order to compare marginal effects of different variables two alternative measures are calculated. Since the marginal effects depend on the value of each variable at which it is obtained, they are not directly comparable. Calculating the centered change over the range (Equation 64) of the variable and also obtaining the centered change of one standard deviation (Equation 65) alleviate this problem.

$$(64)\quad \frac{\Delta \Pr(y_{ni}=1|\bar{x}_n)}{\Delta x_k} = \Pr(y_{ni}=1|\bar{x}_n, Max"x_k") - \Pr(y_{ni}=1|\bar{x}_n, Min"x_k")$$

$$(65)\quad \frac{\Delta \Pr(y_{ni}=1|\bar{x}_n)}{\Delta x_k} = \Pr(y_{ni}=1\left|\bar{x}_n, \bar{x}_k + \frac{s_k}{2}\right.) - \Pr(y_{ni}=1\left|\bar{x}_n, \bar{x}_k - \frac{s_k}{2}\right.)$$

4.4.2 Measures of Fit and Specification Tests

Two measures of fit are reported. Define $L(M_\alpha)$ as the likelihood of restricted model with just the intercept and $L(M_\beta)$as the likelihood of the model with regressors included. The Maddala pseudo-R^2 can be calculated as (Long 1997):

$$(66)\ R^2_{ML} = 1 - \left[\frac{L(M_\alpha)}{L(M_\beta)}\right]^{2/N} = 1 - \exp(-G^2/N) \text{ where } G^2 = -2\ln\left[\frac{L(M_\alpha)}{L(M_\beta)}\right]$$

Another measure of fit is the normed measure of Cragg and Uhler given by:

$$(67)\ R^2_{C\&U} = \frac{R^2_{ML}}{\max R^2_{ML}} = \frac{1 - \left[\frac{L(M_\alpha)}{L(M_\beta)}\right]^{2/N}}{1 - L(M_\alpha)^{2/N}}$$

Since the MLE estimators are asymptotically normal distributed $\hat{\theta} \overset{a}{\sim} N[\theta, I\{\theta\}^{-1}]$, where $\theta = (B_1, B_2, \ldots\ldots B_K)'$, the simple hypothesis $H_0: \beta_k = \beta^*$, β^* being the hypothesized value is carried over using a z-test. If the null hypothesis is true then z is distributed approximately normal with a mean of zero and variance of 1.

$$(68)\ z = \frac{\hat{\beta}_k - \beta^*}{\hat{\sigma}_{\hat{\beta}_k}}$$, where $\hat{\sigma}_{\hat{\beta}_k}$ is the k^{th} element in the diagonal of $Var(\hat{\theta})$.

For a set of restrictions $R\theta = q$ the Wald test is used. W is distributed as chi-square with degrees of freedom equal to the number of constrains. Following Green (2002), the Wald statistic is given by:

$$(69)\ W = (R\hat{\theta} - q)'\{R(Var(\hat{\theta}))R'\}^{-1}(R\hat{\theta} - q)$$

Another statistic used in logit discrete choice models is the Likelihood Ratio test defined as in equation (63). If the null hypothesis is true, then LR is asymptotically distributed as chi-square with degrees of freedom equal to the number of constraints.

$$(70)\quad LR = G^2 = -2\ln\left[\frac{L(M_\alpha)}{L(M_\beta)}\right]$$

5. Survey and Factor Analysis Results

This chapter begins with a description of the demographic characteristics of the surveyed sample followed by the analysis of the choice preferences and purchase behavior reported by the respondents. The last section presents the results of the factor analysis applied to the set of questions used to elicit consumer's behavior regarding the environment and their health.

5.1 Socio-Demographic Information

In total, 300 respondents were recruited to fill out the survey. From them, 280 (93%) valid responses were obtained. The 20 eliminated questionnaires were as follows: fourteen observations (14) were excluded from the analysis because of insufficient information (more than 20% of missing data), five (5) others were excluded because of duplicate information since they were identified as belonging to the same household (one eliminated from each pair) and one (1) was excluded because it was completed by an under aged respondent.

The sample consisted mostly of females (61.4%) with 59% of the sample being below 44 years old. According to the 2002 American Community Survey Profile (ACS) of the U.S. Census Bureau for Multnomah County, the age group with the highest participation was 25-44 with 47% followed by 45-64 with 33%. About 12% of the respondents were aged between 18 and 24 while only 9% of the sample was people aged 65 years or older. The age distribution for people aged 18 or older according to the ACS is similar to that found for this sample. The age group participation for people between 18-24 years old was found by the ACS to be 12%, for the group 25-44 years old 44%, 45-64 31% and 65 and over 13%.

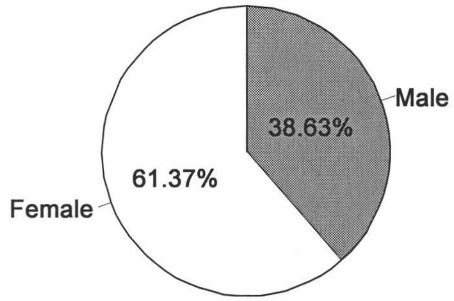


Figure 5.1: Gender Distribution

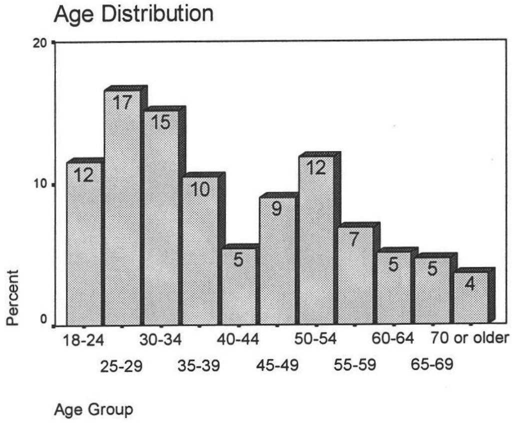


Figure 5.2: Age Distribution

Figure 5.3 illustrates that 49 (17.8%) of the 280 households have an income over $ 100,000, while 16% of the households in the sample have an income less than $20,000. In total, 53% of the respondents have a household income lower than $50,000. The average income reported by the ACS for Multnomah County in 2002 is about $32,000. The average income for this sample (calculating the average for the midpoints of the income categories) is $57,000. The question asking household income in the survey was close ended. The 35% of the interviewed consumers have a 4 year College degree, followed by people having an advanced college degree (29)% (Figure 5.4). Comparing to the ACS results for 2002, the sample used in this research had a higher percentage of people with high educational attainment levels. The ACS results for Multnomah County for 2002 for people 25 years old and over are: Graduate or professional degree (12%), Bachelor's degree (22%), high school (22%) and others (44%).

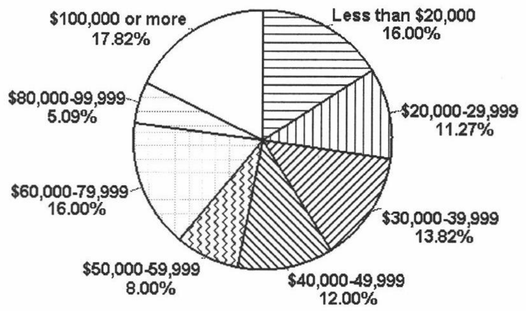


Figure 5.3: Household income level

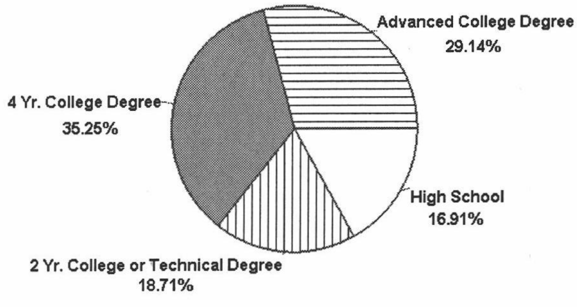


Figure 5.4: Educational level

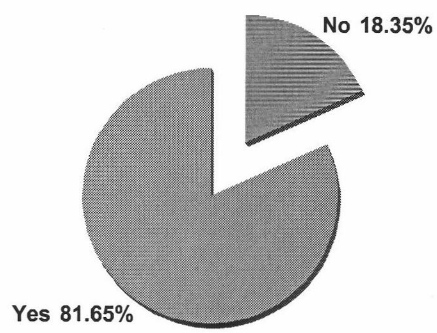


Figure 5.5: Primary Shopper

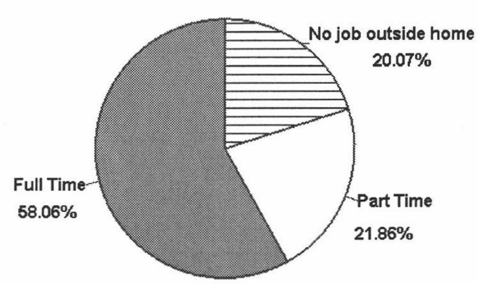


Figure 5.6: Job participation

Corresponding to the gender distribution, almost 82% of the respondents answered as being the primary shopper of the household as seen in Figure 5.5. With respect to job participation, 58% of the consumers answered as having a full time job, 22% as having a part time job, and 20% reported not having a job outside the home (Figure 5.6). Also, 52% of the respondents reported that they owed their current residence while 48% reported as being tenants. Another question asked respondents is if they were a member of an environmental organization. About 27% (76 observations) indicated they are members of this type of organization.

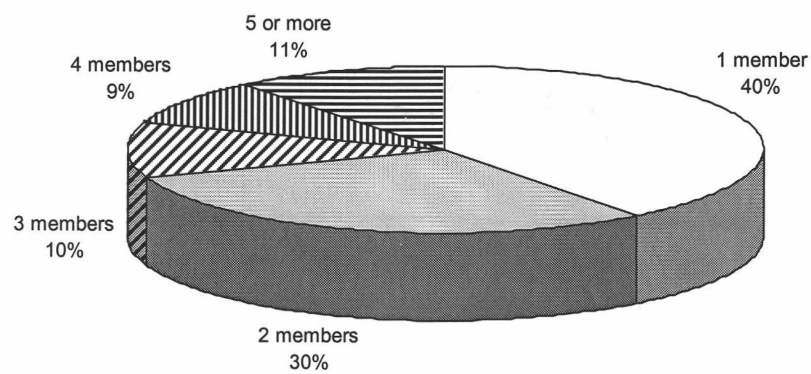


Figure 5.7: Number of household members

Figure 5.7 presents number of members in the respondent's household. The majority of the households have one (40%) or two members (30%), 21% reported three or more individuals. The survey also asked the respondents about their fresh produce shopping behavior. As seen in Figure 5.8, respondents have a similar regularity in shopping fruits and vegetables. The majority of consumers consumed fresh fruits and vegetables one or more times a day. Fewer than 5% of the sample reported having a serving of fruit or vegetables less than once a week.

Table 5.1 shows the participation of each one of the places where consumers reported buying fresh fruits and vegetables. As shown in Table 5.1 about 70% of the survey respondents reported buying fresh produce regularly from supermarkets and groceries, 44.4% reported regular buying at a farmer's market and 38.5% reported buying at natural food stores. The percentages do not sum to 100% since respondents were able to choose two or more store venues.

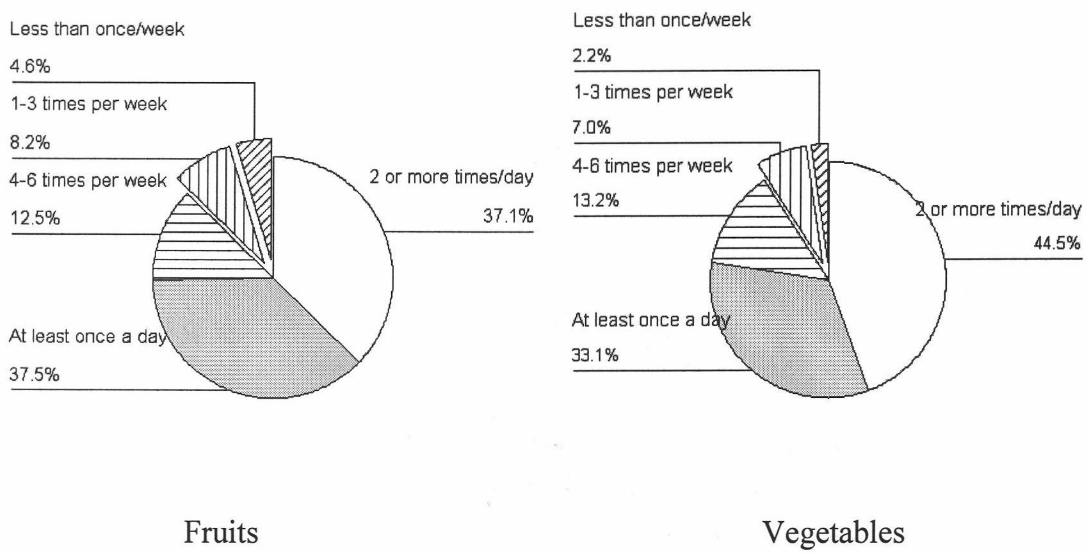


Figure 5.8: How often do you have a serving of fruits and vegetables?

Table 5.1: Where do you regularly buy fresh fruit and/or vegetables?

	Table Responses %
Supermarket/Grocery	70.3%
Farmers Market	44.4%
Natural Food Store	38.4%
Co-operative	31.2%
Directly from Farmer	8.6%
Direct Delivery	.7%

5.2 Product Choice Preferences and Purchases

One of the key aspects of the survey was to collect data about consumer's preferences for vegetables and fruits regarding the choice between organic and conventional produce. Figure 5.9 and Figure 5.10 illustrate respondent's preferences for fruits and vegetables. It is clear once again, the similarity of the results for both products. Given the characteristics of the sample and data collection, where the three locations chosen offer both organic and conventional produce, is not surprising that the vast majority of respondents prefer organic fresh fruits and vegetables in a proportion of 61%. It is important to note that these are not all the same individuals. Conventional shoppers for fruits and vegetables participate in the sample in a proportion of 38%. The other eco-labeled and the don't buy category had less than a 2% response for both products. It is important to recall that the sampling process was deliberately skewed towards target participants that prefer organic. The results on organic preferences in this sample cannot be fully transmitted to the population.

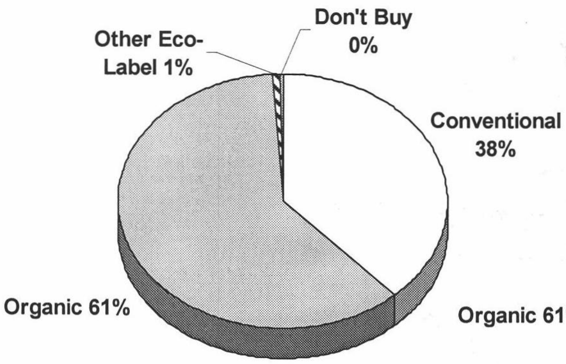


Figure 5.9: Preference choice of Vegetables

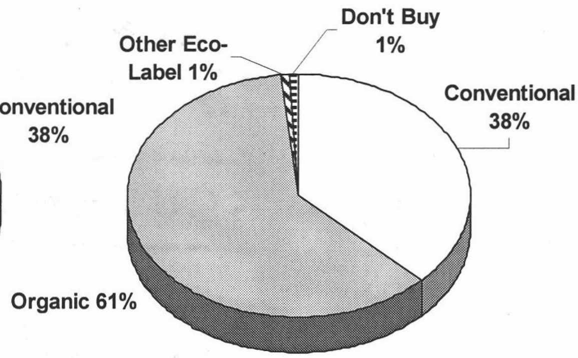


Figure 5.10: Preference choice of Fruits

Consumer's preferences were gathered also for other products. Results for coffee/tea, juice, dairy, eggs, frozen foods, canned foods, meat/poultry and snacks, and cereals were also gathered. It is worth noting that 40% of the respondents reported preferring organic Coffee/Tea, 48% reported preferring organic juice, 39% preferred organic dairy and 42% preferred organic eggs. For canned food; about 26% reported preferring organic and 25% reported preferring organic frozen foods. In the case of meat/poultry and snacks, about 33% of the sample reported preferring organic varieties. These products present a higher percentage of people who prefer conventionally grown products and people who 'do not buy' than the fresh produce categories.

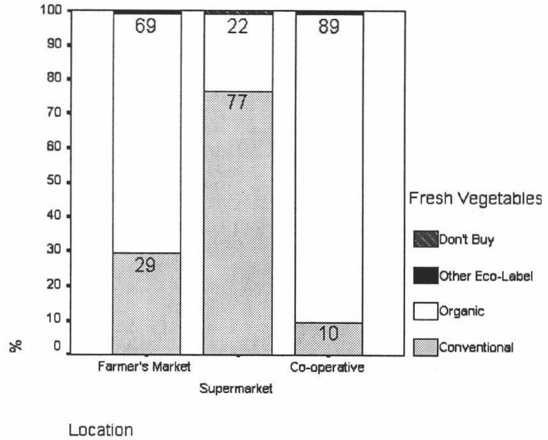


Figure 5.11: Preference choice of vegetables by type of store

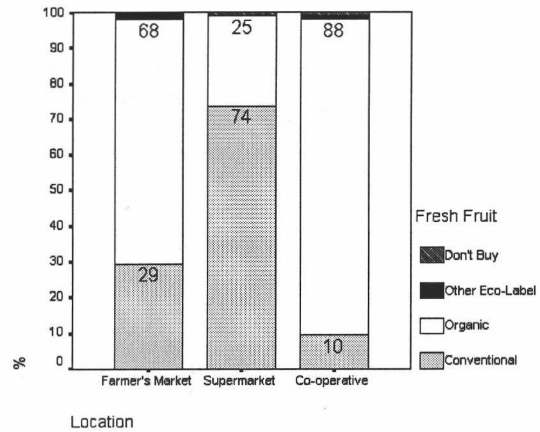


Figure 5.12: Preference choice of fruits by type of store

The cross tabulations of the type of product that consumers preferred by location are presented in Figure 5.11 for vegetables and Figure 5.12 for fruits. Once again the results of the cross tabulations are almost identical for both products. Since the food co-operative is specialized, though not exclusively, in offering organic products, more of its customers prefer organic produce (89% of them for vegetables and 88% for fruits). It is assumed that dedicated organic consumers shop at the co-operative precisely because they offer organic produce. On the other hand, supermarket customers prefer conventionally grown to organic (77% for vegetables and 74% fruits). Even though the supermarket offered organic varieties for most fruits and vegetables, it does not specialize in this market and it attracts a wider class of shoppers, not only dedicated organic buyers. The farmer's market places between the supermarket and the co-operative. About 69% and 68% for vegetables and fruits respectively, prefer organic varieties over conventional ones.

Appendix 3 contains the results of the cross tabulations between demographic variables by type of venue. Analysis of the results showed that respondents at the farmer's market were mostly females (74%) had at least a 4 year college degree (83%), were younger than 44 years old (56%), and had a household income greater than $60,000 (57%). At the supermarket, the proportion between females and males was equal, but 56% of the respondents had at most a 2 year college degree. In this venue it was found the same proportion (50%) between individuals below 44 years and over 44 years old while the majority (54%) reported household income below $60,000. Finally, in the case of the co-operative, 59% of the respondents were female, 64% reported having at least a four college degree and the majority of the individuals (62%) were aged between 25 and 44 years old. 62% of the respondents at the co-operative reported having a household income of $40,000 or less.

In order to compare consumer's preferences for organic produce with actual purchasing behavior, the survey included a question asking respondents what percentage of fresh fruits and vegetables purchases are organic. Figure 5.13 and Figure 5.14 illustrates the frequencies for organic purchases by percentage category for vegetables and fruits respectively. The two categories with the greatest number of responses for both products are 1-10% and 91-100%, each with 51 observations. From the 51 respondents for the 1-10% level, cross tabulations show that 43 (84%) for fruits and 45(90%) for vegetables preferred conventional. In the case of the 91-100% category, cross tabulations revealed that 96% of the people preferred organic fruit while 98% preferred organic vegetables. Given the sampling process, these results do not hold for the population since organic consumers were deliberately over sampled to collect more information about this specific group.

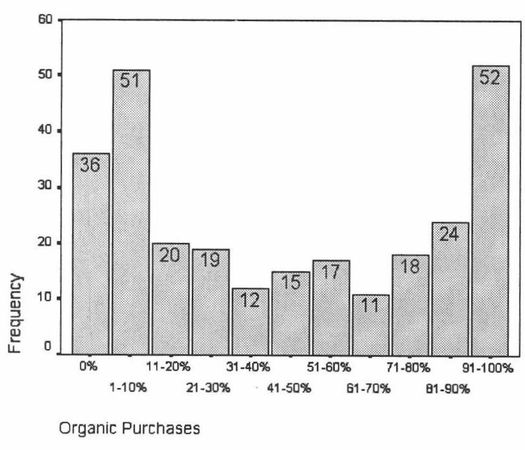


Figure 5.13: Frequencies for Organic Vegetable Purchases

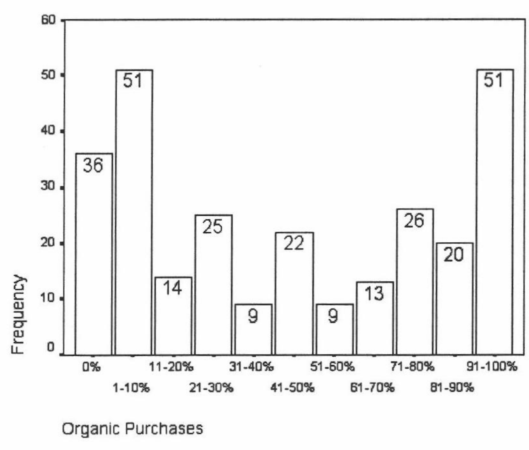


Figure 5.14: Frequencies for Organic Fruit Purchases

The survey also contained three questions about willingness to pay a premium for organic certified fresh produce. One of the questions asked how much the person would be willing to pay for a pound of organic apples if the price of conventional is

$0.99 per pound. Six prices were analyzed, one lower than conventional ($0.79/pound), the same ($0.99/pound) and four higher than the price of conventional apples. The majority of the respondents, 102 (36.8%) were willing to pay a 30 cent price premium, 63 (22.7%) were willing to pay a 60 cent premium while 61 (22%) were not willing to pay any premium. From the respondents willing to pay a premium for organic apples, 34 answered preferring conventional fruits, 64 organic fruits and 2 that preferred eco-label apples. Figure 5.15 shows the cumulative frequencies for individuals willing to pay a premium for organic apples by venue. More consumers at the co-operative and farmer's market were willing to pay higher premiums for organic apples than individuals surveyed at the supermarket.

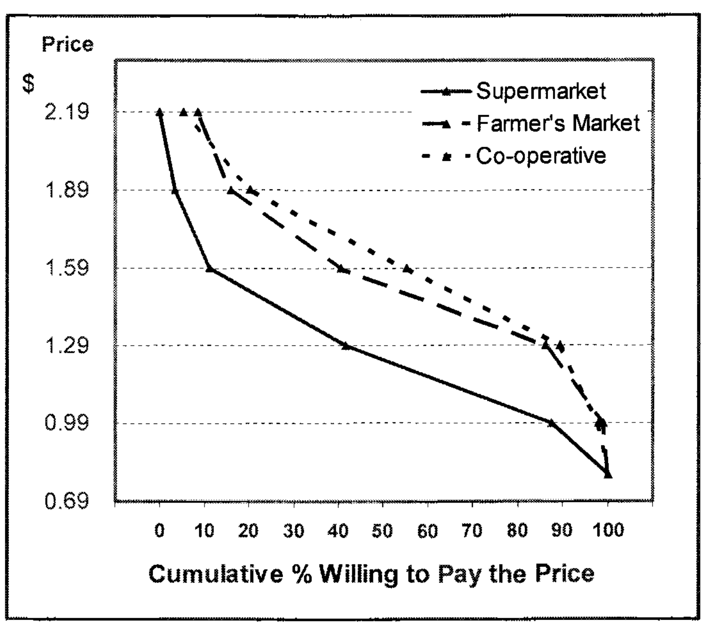


Figure 5.15: How much would you pay for a pound of organic apples if the price of conventional apples is $0.99 / pound?

People who were not willing to pay a premium for organic apples corresponded mostly (74%) to people who preferred conventional fruits. Not surprisingly, cross tabulations between this question and the choice of fruits revealed that most of the people willing to pay a premium of more than 30 cents were people who preferred to buy organic fruit; the reverse was found for people who were willing to pay a lower price for organic apples.

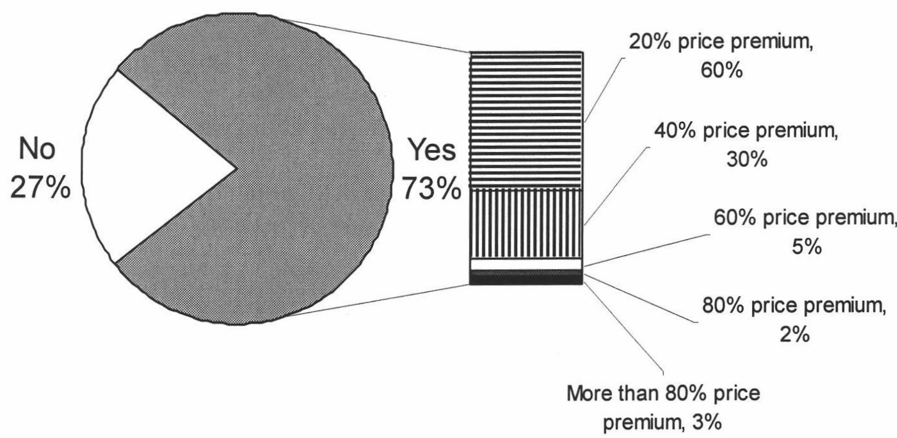


Figure 5.16: Are you willing to pay a premium for organic vegetables?

Two other questions were formulated regarding willingness to pay a premium for certified organic vegetables and fruits. Tabulation results for these two products revealed similar results. Figure 5.16 presents the results for vegetables only. 73% of the respondents were willing to pay a premium for organic vegetables and 72% for fruits. From the people who answered yes, 60% were willing to pay a 20% price premium for vegetables (61% for fruits), 30% a 40% premium for vegetables (29% for fruits); while consumers willing to pay more than a 60% premium were less than 10%. Cross tabulations with consumer preferences revealed some discrepancy in that

24% of consumers who were willing to pay a premium for organic vegetables (22% fruits) preferred conventional.

In order to determine how important the organically grown attribute is compared to others such as price or taste, the survey included a question in which respondents were able to rate eleven attributes from "Not important" to "Very important" for fresh produce and also rank[8] the top 5 of them. Table 5.2 shows the weighted rankings[9] of the eleven attributes classified by choice of fresh produce[10]. The top five attributes that people who preferred conventional produce referred when shopping were taste, price, appearance, ripeness and nutritional value. The "organically grown" attribute ranked last, even lower than "pesticide use reduction". It is clear that consumers from the sample who preferred conventional produce are driven by product attributes such as taste, price or appearance.

Table 5.2: Ranking of fresh produce attributes by choice preferences

Conventional Fresh Produce			Organic Fresh Produce		
Ranking	Attribute	Weighted Ranking	Ranking	Attribute	Weighted Ranking
1	**Taste**	**884**	**1**	**Organically Grown**	**1304**
2	**Price**	**764**	**2**	**Taste**	**1277**
3	**Appearance**	**742**	**3**	**Free of GMO**	**1071**
4	**Ripeness**	**601**	**4**	**Pesticide Use Reduction**	**1019**
5	**Nutritional Value**	**510**	**5**	**Nutritional Value**	**935**
6	In season	405	6	Locally Grown	930
7	Pesticide Use Reduction	334	7	Price	765
8	Free Of GMO	332	8	In season	686
9	Locally Grown	325	9	Appearance	662
10	Size	284	10	Ripeness	661
11	**Organically Grown**	**195**	11	Size	264

[8] Since the design of this question was close-ended (as all the survey was) it presented some confusion to many of the respondents who tended to rank all the attributes not the top five only. These observations were dropped for the ranking.
[9] The weighted rankings are simply the total sum of the product between the importance (very important=3, moderately important=2, not important=1) and the ranking received by each attribute in a scale from 1 to 5, five being the highest ranking.
[10] The question referred to both fruits and vegetables.

On the other hand, consumers who preferred organic fresh produce ranked the organically grown attribute over price, appearance and even taste. The top five attributes for this group are organically grown, followed by taste, free of GMO, pesticide use reduction and nutritional value. These results suggest, given the sample, that organic shoppers are less price oriented and more interested in production characteristics.

Another question included in the survey sought to determine respondent level of awareness of four health and environmental issues as it influenced buying fruits and vegetables. Question 18 asked the respondents to rate their level of concern with respect to: a) Pesticides residues on food; b) Preservatives and additives; c) Presence of chemicals that cause cancer and d) Possible effects of GMO products on your health[11].

It is very clear that consumers who preferred organic produce are more concerned with these issues than their conventional counterparts. Sixty seven percent (67%) of organic buyers are very concerned about "pesticides residues on food", more than three times as many as conventional buyers (20%). The same pattern is true for "preservatives and additives" where 59% of them where very concerned with this issue, versus only 17% in the case of people who preferred conventional. 76% of organic shoppers answered as being very concerned with "presence of chemicals that cause cancer" while only 39% of conventional buyers answered in the same way. Finally, for the question related to GMO products (Genetically Modified Organisms)

[11] The cross tabulation of Question 18 with choice of fruits yielded almost identical results.

67% of organic and 17% of people who preferred conventional were very concerned with the possible effects of GMO products on health.

Question 19 and 20 were designed to measure respondent's level of concern towards environmental issues. These two questions were designed as scales from 1 to 10 where the respondent had to choose the relative importance when comparing the issues. Question 19 asked, "when purchasing foods, what is the importance of buying products friendly to the environment versus lower costs of food?"; with 1 meaning buying eco friendly products is all important and 10 meaning lower costs of food is all important.

Figure 5.17 illustrates the histograms for the results of this question by choice of vegetables[12]. Clearly, conventional shoppers are distributed around the midpoint of the scale with a mean of 5.6. On the other hand, organic shoppers were concentrated at the left tail of the distribution with 120 observations located between 1 and 4 and a mean of 3.6. This result implies that organic buyers are more concerned with protecting the environment when buying produce than conventional shoppers.

Question 20 examined "reducing harmful chemicals" versus "lower costs of food". Figure 5.18 shows the histograms for this question classified by choice of vegetables. Clearly, organic shoppers are more concerned about reducing harmful chemicals when buying food than conventional ones. The mean for organic buyers is 2.26 while the mean in the case of conventional buyers is 4.9. It is worth noting that most organic buyers positioned themselves between 1 and 2, while conventional shoppers had a distribution concentrated around the mean.

[12] The results for fruits and vegetables were almost identical, only results for the later are presented here.

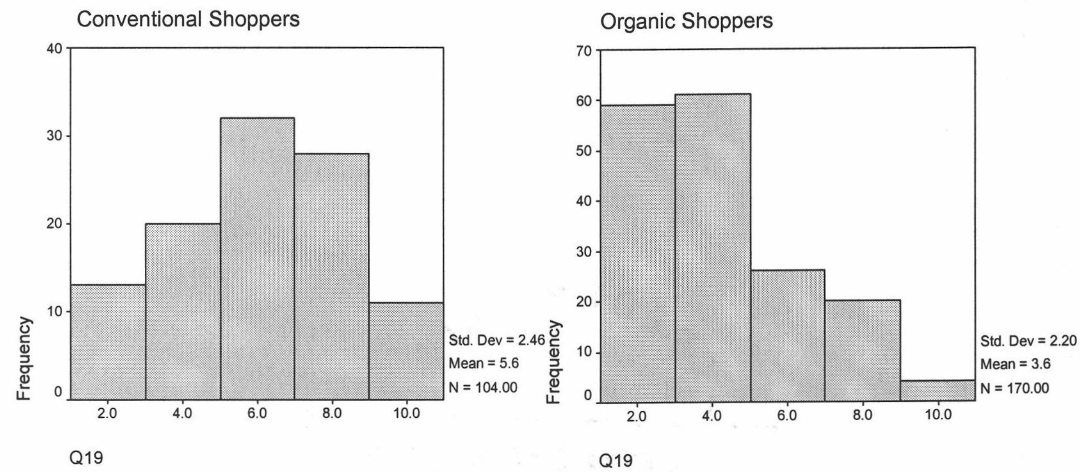


Figure 5.17: Histograms for the importance of buying products friendly to the environment versus lower cost of food.[a]

a= On a scale 1 to 10, where 1 means buying products friendly to the environment and 10 means lower costs of food is all important.

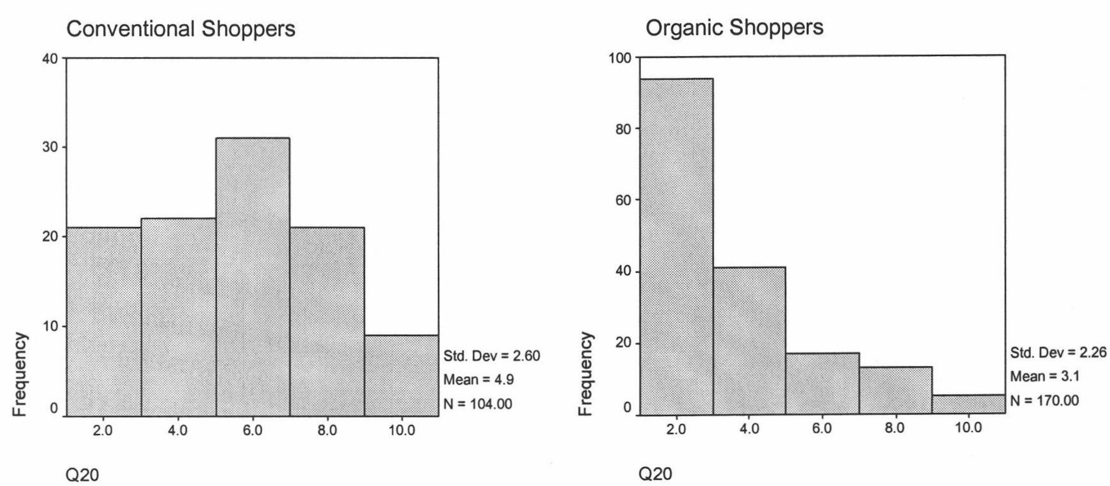


Figure 5.18: Histograms for the importance of reducing harmful chemicals versus lower cost of food[a].

a= On a scale 1 to 10, where 1 means reducing harmful chemicals and 10 means lower costs of food is all important.

5.3 Socio Demographic Information by Choice of Fresh Produce

This section presents the cross tabulation of demographic information by choice of vegetables. Since the results for fruits are similar to those for vegetables, they are omitted in this section. Figure 5.19 presents the results of gender by preference choice of vegetables. Females have a slightly bigger (64.3%) participation in the group of organic shoppers than in the conventional group (56.6%). In the case of level of education, Figure 5.20 shows no noticeable difference between the distributions among both groups. Organic shoppers have a higher participation of individuals with advanced college degree (32.8%) than conventional (24%) and lower in the case of high school and 2 year college degree. Professionals with a four year college degree have about the same participation of 36% in the organic and conventional group.

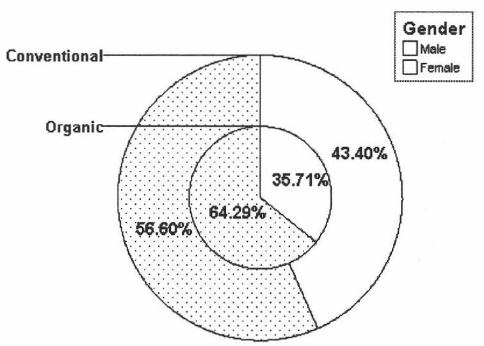


Figure 5.19: Gender by choice of vegetables

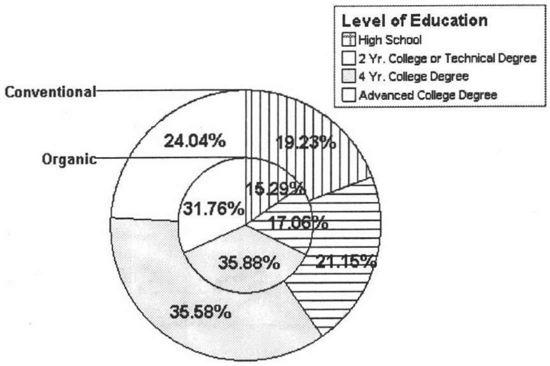


Figure 5.20: Level of Education by choice of vegetables

Even though there are no noticeable differences between organic and conventional shoppers in the case of gender and level of education, the age distribution seen in Figure 5.21 presents clear differences. There are more people in the younger age groups in the case of organic shoppers than in the case of conventional. About 69% of organic shoppers are age between 18 and 44, while in comparison 45% of conventional shoppers are in the same age group. The group with highest participation in the case of organic is 25-29 years with 20%, while in the case of conventional 15% are aged between 50 and 64 years.

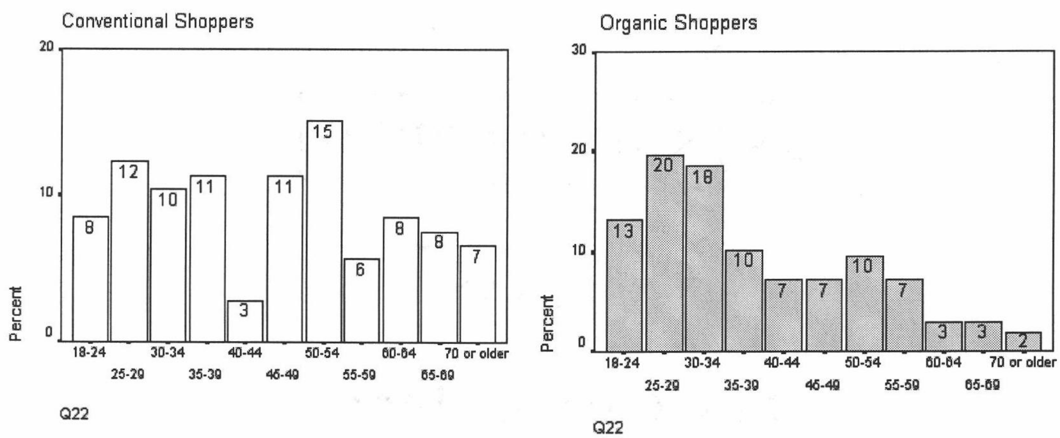


Figure 5.21: Age distribution by choice of vegetables

Another demographic characteristic that presents differences between organic and conventional buyers is household income. About 60% of organic shoppers reported a household income less than $50,000, in comparison with 42% of the conventional group (Figure 5.22).

Analysis of the composition of the household by number of members revealed more single member households among organic buyers with 45% of people who preferred organic and 32% for conventional shoppers living in one member households. Analysis of the member of households by age tabulated by choice of vegetables yielded similar results, except for the age group from 25-39 years which was in greater proportion composed by single individuals for organic, as a result of the greater proportion of one member households explained above.

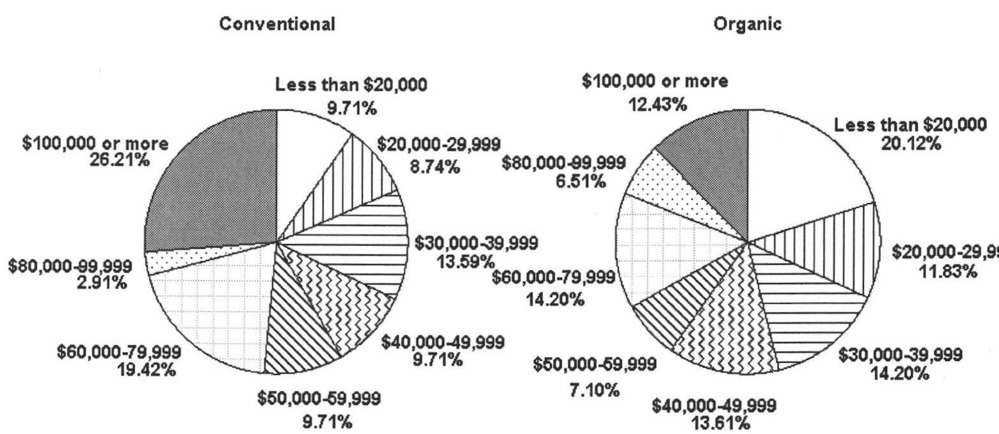


Figure 5.22: Income Distribution by choice of vegetables

There was no significant difference between organic and conventional buyers as to whether respondent was household's primary shopper. On the other hand 57% of organic shoppers rented in the sample their current residence in comparison to 36% of conventional buyers. This difference appears to be related to age. Moreover, only 11% of the conventional buyers in the sample reported being members of an environmental organization compared to 38% of the organic buyers. With respect to occupation, about 62% of individuals who preferred organic have a full time job in comparison with 53% for conventional. Finally, there are more organic buyers who grew up in a Urban-Rural mix area than conventional buyers, but the relative participation for only urban or rural areas are very similar.

5.4 Factor Analysis Results

As explained in section 4.3, respondents were questioned about their behavior towards the environment and their attitude and behavior towards their health. Questions were asked on recycling, energy conservation, how they select goods with respect to their effect on the environment, their physical activity and attitude towards personal health responsibility. Respondents had to indicate their level of agreement with the 29 behavioral/attitudinal questions on a 5 point scale ranging from "never true" to "always true". Following the methodology presented in Chapter 4, this section presents the results of applying factor analysis to these questions.

As discussed in section 4.3, factor analysis attempts to identify underlying variables (factors or components[13]) that explain the pattern of correlations within a set of observed variables. Factor analysis is designed to measure latent variables and facilitates a reduction in the number of variables incorporated into regression models by identifying the principal relations that exist between those variables. Factor Analysis was carried out using the principal components approach and a Varimax rotation. The factor analysis can be applied to the health and environmental questions

[13] The underlying variables are typically called "factors" in Factor Analysis. When using the Principal Components approach, these factors are referred to as "components".

separately or jointly. Both approaches were examined and produced the same factors. The joint results were used as will be discussed below[14].

The factor analysis from the set of environmental and health questions detected the presence of 7 components. The environmental questions produced two principal components and four components were obtained from the health related questions. Table 5.3 presents the initial eigenvalues, the extraction sums of squared loadings and the rotation sums of squares. The first seven components have eigenvalues greater than one with the last having an eigenvalue very close to 1. Since an eigenvalue is the ratio of the variance among the data accounted for/explained by a component compared with the amount that would be accounted for by chance, only factors with eigenvalues greater than one explain more variance than would be expected.

Together, the seven factors identified explained 64 % of the total variance observed in the 29 questions. Component 1 explains 30.2% in the unrotated solution (Extraction sums of Square Loadings) while component 2 explains 9.3%. After the factor matrix was rotated, component 1 explains 19.9% while component 2 explains 10.6% of the total variance (Rotation Sums of Square Loadings). The rest of components each explain less than 5% of the total variance. Since the rotated matrix is just a product of the factor loadings by an orthogonal matrix, the total variance explained together by the seven components remains at 64%. Rotation is undertaken to facilitate the identification of the components by presenting stronger factor loadings for those variables that are highly correlated with the components. The scree

[14] Factor Analysis was carried on using the Statistical Package for Social Sciences (SPSS) version 11.5

test, plotting the eigenvalues of the components, also confirms the presence of seven principal components in this set of variables (Appendix 6). Again there are seven eigenvalues greater than one after which the plot becomes flatter.

Table 5.3: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.766	30.228	30.228	8.766	30.228	30.228	5.788	19.957	19.957
2	2.715	9.363	39.591	2.715	9.363	39.591	3.067	10.575	30.532
3	1.913	6.598	46.189	1.913	6.598	46.189	2.445	8.432	38.965
4	1.723	5.940	52.130	1.723	5.940	52.130	2.203	7.597	46.562
5	1.289	4.445	56.574	1.289	4.445	56.574	2.066	7.123	53.685
6	1.163	4.010	60.584	1.163	4.010	60.584	1.677	5.782	59.467
7	1.012	3.488	64.072	1.012	3.488	64.072	1.335	4.605	64.072
8	.877	3.026	67.098						
9	.848	2.924	70.021						
10	.785	2.709	72.730						
11	.722	2.489	75.219						
12	.671	2.315	77.534						
13	.641	2.210	79.744						
14	.588	2.029	81.773						
15	.581	2.002	83.774						
16	.534	1.843	85.617						
17	.500	1.724	87.341						
18	.447	1.542	88.883						
19	.420	1.447	90.330						
20	.382	1.317	91.646						
21	.371	1.278	92.924						
22	.362	1.248	94.172						
23	.304	1.048	95.220						
24	.295	1.016	96.235						
25	.249	.857	97.093						
26	.245	.846	97.939						
27	.224	.772	98.711						
28	.215	.742	99.453						
29	.159	.547	100.000						

Extraction Method: Principal Component Analysis.

Table 5.4 Rotated Component Matrix(a)

Component Description	Variable	Component 1	2	3	4	5	6	7
Environmental Purchasing Behavior (EPB)	c) I buy environmental friendly products	0.829	0.189	0.079	0.033	0.072	-0.001	0.064
	g) I have switched product for environmental reasons	0.778	0.205	0.299	0.030	0.032	-0.008	0.150
	i) I will buy from companies that are ecologically irresponsible	0.762	0.070	0.096	0.017	0.130	-0.011	0.160
	h) I have convinced others not to buy products harmful to the environment	0.758	0.086	0.241	0.038	0.116	0.009	0.112
	k) I have purchased products because they cause less pollution	0.724	0.194	0.421	0.012	0.010	0.047	0.022
	m) I do not buy household products that harm the environment	0.676	0.249	0.353	0.084	0.055	0.121	0.091
	l) I try only to buy products that can be recycled	0.666	0.224	0.470	0.139	-0.055	0.040	-0.057
	b) Do you look for the dolphin-safe label when you buy tuna?	0.656	-0.016	-0.085	0.062	0.184	0.035	-0.086
	Free of GMO important when buying fruits and vegetables	0.615	-0.086	-0.028	0.105	0.437	-0.066	-0.104
	n) I avoid foods containing nitrites or preservatives	0.565	0.446	-0.028	0.148	0.169	0.016	0.062
Nutrition Management (Nutrition)	p) My daily diet is nutritionally balanced	0.206	0.795	0.083	0.199	-0.063	-0.060	0.178
	q) I eat 5 servings of fruits and vegetables every day	0.199	0.710	-0.022	0.304	-0.064	-0.033	0.099
	o) I am interested in information about my health	0.310	0.639	0.127	-0.033	0.235	-0.116	-0.167
	g) I try to avoid high levels of cholesterol in my diet	0.029	0.528	0.144	0.135	0.282	0.228	-0.002
	d) Good health takes active participation on my part	0.056	0.487	0.009	0.156	0.399	-0.131	0.029
	a) I read more health related articles than I did 3 years ago	0.033	0.475	0.301	0.081	0.402	0.095	-0.292
Energy Conservation and Recycling Behavior (ECRB)	e) I buy energy efficient light bulbs	0.073	-0.064	0.677	0.004	0.164	-0.027	0.438
	j) I have tried very hard to reduce electricity use	0.234	0.161	0.607	0.232	0.184	-0.037	-0.086
	f) I purchase recycled paper	0.475	0.056	0.570	0.091	0.025	-0.033	0.275
	a) I recycle paper, cans or bottles	0.300	0.085	0.560	0.082	-0.178	0.012	0.045
Physical Fitness (Fitness)	t) I try to exercise at least 30 min a day, 3 days a week	0.114	0.288	0.002	0.814	0.044	-0.050	0.034
	r) I regularly participate in outdoor activities	0.115	0.285	0.075	0.780	0.014	-0.132	0.182
	h) I exercise more than I did 3 years ago	0.001	-0.015	0.235	0.695	0.163	0.222	-0.170
	s) I try to avoid stressful situations	0.061	0.217	0.208	0.337	0.251	0.057	0.260
Health Environment Sensitivity (HEResp)	c) I'm concerned about my drinking water quality	0.197	0.085	-0.014	0.110	0.801	0.054	0.139
	b) I worry that there are harmful chemicals in my food	0.350	0.260	0.080	0.024	0.692	0.008	-0.081
Personal Health Responsibility (PHResp)	e) It is the doctor's job to keep me well	0.037	0.011	0.034	0.014	0.022	0.871	-0.018
	f) My health is outside my control	0.029	-0.062	-0.063	-0.003	0.001	0.837	0.131
	d) I use rechargeable batteries[b]	0.182	0.066	0.184	0.064	0.000	0.131	0.794

The next step is to identify and determine the characteristics of the components. Table 5.4 presents the rotated component matrix with the loadings for each component and variable. The shaded cells show the variables that are highly correlated to the components. It can be deduced that the variables with the highest loadings on the first component express some type of purchasing behavior. In fact, the variable with highest loading refers to "I buy environmental friendly products" with 0.83. The rest of the variables also indicate purchasing behavior associated with reducing environmental harm (pollution, dolphin-safe, etc.). Because of these relationships Component 1 is called "Environmental Purchasing Behavior" (EPB), similar to what Straughan and Roberts (1999) called Environmental Conscious Consumer Behavior, which "measures the extent to which individual respondents purchase goods and services to have a more positive (less negative) impact on the environment".

The second component (10.6% of variance explained) was highly correlated with two variables "My daily diet is nutritionally balanced" and "I eat 5 servings of fruits and vegetables every day" with loadings of 0.79 and 0.71 respectively. This component was also moderately correlated with, "I am interested in information about my health" (0.64) and "I try to avoid high levels of cholesterol in my diet" (0.53). Two other variables were correlated "Good health takes active participation on my part" and "I read more health related articles than I did 3 years ago". The last two variables do not have much influence on the interpretation since their coefficients are less than 0.5 and were also correlated with the fifth component. Following the interpretation of Kraft and Goodell, this factor can be designated as "Nutrition

Management Behavior" (Nutrition). This factor expresses respondent's concern and behavior regarding their nutrition, having a nutritionally balanced diet, eating vegetables and fruits every day and avoiding nitrites or preservatives. It is interesting to note that this component is highly correlated with a variable expressing explicit consumption of fruits and vegetables, as part of consumer's nutrition management.

The third component is correlated with four variables. "I buy energy efficient light bulbs" has the highest loading (0.677) followed by "I have tried very hard to reduce electricity use" with 0.60." The other two correlated variables are "I purchase recycled paper" and "I recycle paper, cans or bottles". Hence, component 3 has been referred to as "Energy conservation and recycling behavior" (ECRB). This component is related with behavior towards energy conservation and recycling rather than purchasing behavior as the first component.

The fourth component was named "Physical fitness" (Fitness). It involves items involving exercise and sports. It represents "action-oriented behaviors related to improving and maintaining one's physical condition" (Kraft and Goodell, 1993). The variable with the highest factor loading with this component is "I try to exercise at least 30 min a day, 3 days a week" with 0.814. "I regularly participate in outdoor activities" and "I exercise more than I did 3 years ago" had loadings of 0.78 and 0.69 respectively.

The fifth component has a very clear interpretation with correlations on two variables: "I'm concerned about my drinking water quality" and "I worry that there are harmful chemicals in my food". This component was interpreted as "Health environment sensitivity" in the Kraft and Goodell study, which "involves a clear concern for the impact of the environment on one's health."

The sixth component, "Personal Health Responsibility" (PHResp), reflects an individual's willingness to accept personal responsibility for his/her own health. Two variables were highly correlated with this factor, "It is the doctor's job to keep me well" and "My health is outside my control" with loadings of 0.871 and 0.837 respectively. About 80% of the respondents either disagree or strongly disagree with these two statements (Appendix 5).

The seventh and final factor corresponds to only one variable "I use rechargeable batteries". This variable was expected to be included in the energy conservation and recycling behavior factor and seems to be more a sample specific result. Moreover, the eigenvalue of this variable is 1.012 (almost one) and the variance that this variable contributes to the factor analysis is less than 5%. Since this variable does not add significant insights into the factor analysis interpretation, it is reported here but it was not considered in the final model.

It is important to note that environmental purchasing behavior was found correlated with the "Free of GMO" variable that expresses the importance of having products free of genetically modified organisms when buying fruits and vegetables. This question was not included in either of the original set of questions but was integrated into the analysis given the importance of GMO issues in the organic public debate. It is also worth noting that while the loading between this variable and EPB is 0.62, it also exhibits a moderate correlation with the Health Environmental Sensitivity factor (0.44).

In summary, two components were found related to environmental motives (EPB and ECRB), four factors related to health motives (Nutrition, Fitness, HESens and PHResp). All of the 13 environmental questions, except for "I use rechargeable batteries", were correlated with one of the two environmental factors. All fifteen of the health variables were correlated with the four health factors, exactly as in the case of the Kraft and Goodell (1993) study. As noted earlier the intention was to use more behavior rather than attitude and believe measuring variables in this analysis. The questions identifying components 5 and 6 and the GMO importance are exceptions to this intention.

The scores for each component and individual were calculated as explained in Chapter 4. Six factor scores were calculated, two for the environmental components and four for the health components. The seventh component was excluded from the analysis as explained above. These factors were used as independent variables in a logit regression as explained in the next chapter.

6. Results

This Chapter reviews the empirical specification used to model consumer's preferences and purchases for organic produce and the results obtained from applying the model to the data collected using the survey. The empirical specification is derived according to the general empirical model of Section 3.3 and solved using a multinomial logit as explained in Section 4.4. Two discrete choice models are considered in the analysis. The first examines consumer's preferences for organic produce and the second purchases of organic or conventional produce.

The first section discusses the empirical specification and describes the variables used as regressors in the discrete choice models. The subsequent four sections pertain to the regression results obtained from the empirical models. The results are analyzed and interpreted for consumer's preference and purchasing models. The chapter ends with a brief review of alternative models and functional forms tested.

6.1 Empirical Specification and Variable Description

In Chapter 3 Random Utility Theory was used to develop a model for the choice of organic produce, and the suitability of the Multinomial Logit model to estimate it was demonstrated. From this framework two discrete choice models were designed.

Equation (71) defines a discrete choice model for preferences of organic produce while equation (71') defines the choice of organic produce in terms of purchase levels. The second model (purchases) was broken into two categories, greater than and less than 50% of organic purchases.

$$(71)\ y_{ni} = \begin{cases} 1 & \text{if individual } n \text{ prefer organic produce (i=1), } U_{n1} \geq U_{n2} \\ 0 & \text{if individual } n \text{ prefer conventional produce (i=2), } U_{n1} \leq U_{n2} \end{cases}$$

and also as:

$$(71')\ y_{ni} = \begin{cases} 1 & \text{if individual } n \text{ purchases organic produce in a proportion greater than 50\% (i=1), } U_{n1} \geq U_{n2} \\ 0 & \text{if individual } n \text{ purchases organic produce in a proportion less than 50\% (i=2), } U_{n1} \leq U_{n2} \end{cases}$$

While the purchase model could be estimated with a continuous dependent variable based on each 10% buying level, this analysis continues with the binomial discrete choice model for two reasons. First, because this makes it comparable to the preference model and second because the distribution of percentage of organic purchases does not conform to assumptions for linear estimation. As discussed in section 5.2 the distribution of the frequencies for organic produce were greater for two particular groups: 1-10% and 91-100%. The U-shaped distribution observed in Figure 5.13 for vegetables and Figure 5.14 for fruits, supports breaking the categories into two groups. The breakpoint chosen was 50% and it is somewhat arbitrary.

Other alternatives such as using three or more categories under a multinomial logit formulation were also prohibited due to sample conditions. Under discrete choice models, and particularly logit and probit models, it is necessary to have sufficient variation in the values taken by the independent variables for each category of the dependent variable. If this condition is not met, the estimation procedure can breakdown.

The probability that the utility obtained from organic produce is higher than the utility from conventional produce can be derived using equation (21).

$$(21)\quad \Pr(Y_n = 1) = \Pr\left[U_{n1} \geq U_{n2}\right]$$
$$= \Pr\left[x_n \mathrm{B}_1 + \varepsilon_{n1} \geq x_n \mathrm{B}_2 + \varepsilon_{n2}\right]$$
$$= \Pr\left[(x_n \mathrm{B}_2 + \varepsilon_{n2}) - (x_n \mathrm{B}_1 + \varepsilon_{n1}) \leq 0\right]$$

By the use of equation (21) and (47) the Conditional Probability of choosing alternative i^{th} given the independent variables (x_n) is obtained for a linear-in-parameters logit as it is defined by equation (49') for J=2:

$$(49')\quad \Pr(Y_n = i|x_n) = \frac{\exp(x_n \mathrm{B}_i)}{\sum_{j=1}^{2} \exp(x_n \mathrm{B}_j)},\quad i,j\text{=1,2 ; B}_1\text{=0}$$

The final empirical specification of the probability of choosing organic under the multinomial logit model has the following form:

$$\begin{aligned}(72)\quad \Pr ob(y_{ni} = 1) = F(&\beta_0 + \beta_1 \Pr imary_i + \beta_2 Female_i + \beta_3 Age_i + \beta_4 Kids_i \\ &+ \beta_5 Education_i + \beta_6 Income_i + \beta_7 EPB_i + \beta_8 Nutrition_i + \beta_9 ECRB_i \\ &+ \beta_{10} Fitness_i + \beta_{11} HESens_i + \beta_{12} PH\mathrm{Re}sp_i + \beta_{13} Member_i + \beta_{14} \Pr ice_i \\ &+ \beta_{15} Appearance_i + \varepsilon_i)\end{aligned}$$

Where F corresponds to the CDF of the logistic distribution. These models are estimated for fresh fruits and for fresh vegetables individually. For the preference model the dependent variable is a binomial variable coded as 1 if individuals prefer organic vegetables or organic fruits, 0 if consumers prefer conventional products.

Sixty-two percent of the sample preferred organic fresh vegetables and organic fresh fruit. For the purchase model the dependent variable is coded as 1 when 51-100% of the vegetables or fruit purchases are organic and 0 otherwise, but only those individuals who reported preferring are included. Respectively, 70% and 68% of this sub-sample reported buying organic vegetables and fruits in a proportion greater than 50%. These percentages are expected to be smaller for a general population.

The descriptive statistics for the regressors are presented in Table 6.1 . Fifteen variables were used in the final models as explanatory variables[15]. Five were dummy variables, and ten were scalar or continuous. "Primary" was a discrete variable taking the value of one if the individual was the primary shopper of the house, 0 otherwise; 82% of the sample reported being the primary shopper. "Female" was coded as one if the respondent was female (61% of the sample). "Age" represents the respondent's age in years, and individual ages were based in the midpoint of the age category in which respondents placed themselves. About 12% of the individuals belonged to the lowest age group (18-24) while about 4% were aged 70 or above.

Another discrete regressor is "Kids" which was coded as one if the respondent reported the presence of individuals less than 18 years old in the household (23% of the sample). "Education" represents the highest level of education attained by the individual. This variable is coded as one if the education level is high school, 2 if the individual attended a two year college, 3 if the individual had a four year college and

[15] Other variables that were tested but were not found significant and were excluded from the final model included: having a Job, area of growing up, number of members in the household, owning the current residence and having illnesses in the past.

4 if he/she has an advanced college degree (MS, PhD, MD, JD). A number of variations on these variables were examined and will be discussed in section 6.5.

Table 6.1: Descriptive statistics for regressors

	Mean	Std.Dev.	Minimum	Maximum	Description
Primary	0.82	0.39	0	1	1 if individual is primary shopper of the household; else 0
Female	0.61	0.49	0	1	1 if individual is female; else 0
Age	41.0	14.7	21.0	72.0	Individual's age in years
Kids	0.23	0.42	0	1	1 if there are kids less than 17 years old in the household; else 0
Education	2.76	1.05	1	4	1 if the individual has been in high school, 2 in a 2 year college, 3 in a four year college, 4 if advanced degree
Income	5.69	3.56	1.5	12.0	Household income in tens of thousands of $
EPB	0.0	1.0	-2.97	1.79	Factor Scores for Environmental Purchasing Behavior (EPB)
Nutrition	0.0	1.0	-4.57	2.49	Factor Scores for Nutrition Management
ECRB	0.0	1.0	-3.86	2.66	Factor Scores for Energy Conservation and Recycling Behavior (ECRB)
Fitness	0.0	1.0	-3.79	2.02	Factor Scores for Physical Fitness
HESens	0.0	1.0	-3.66	2.02	Factor Scores for Health Environment Sensitivity
PHResp	0.0	1.0	-1.85	3.89	Factor Scores for Personal Health Responsibility
Member	0.28	0.45	0.0	1.0	1 if individual is member of an environmental organization, 0 otherwise
Price	2.38	0.58	1.0	3.0	1 if individual thinks price is not important, 2 if it is moderately important and 3 if price is very important when buying fruits or vegetables
Appearance	2.50	0.59	1	3	1 if individual thinks appearance is not important, 2 if it is moderately important and 3 if appearance is very important when buying fruits or vegetables

For regression purposes, "Income" taken on a multiple choice range was assigned its midpoint in 10s of thousands of $. Incomes greater than $100,000 were assigned a 1.2 and those under $20,000 a 1.5. The mean income observed in the sample was 5.69. "Member" is a dummy variable coded as one if the individual reported being member of an environmental organization, 0 otherwise: 28% of the respondents in the sample were members of an environmental organization.

The next six regressors are the factor scores obtained from the factor analysis reported in Chapter 5. The factor scores are simply a calculation from the measures of the 6 factors identified for each individual ($\hat{F} = Y_c\hat{B}_1 = Y_cS^{-1}\hat{\Lambda}$ - See Section 4.3.2). These six factor scores have a mean of 0 and standard deviation of one as result of the assumptions made in the factor analysis. Larger factor scores represent individuals that showed greater commitment to the issues represented by each component obtained in the factor analysis. The opposite is also true. For example, scores for EPB (Environmental purchasing behavior) range from -2.97 to 1.79, the maximum possible value observed (1.79), would correspond to an individual who answered "always true" to all the questions related with this factor. The same logic applies to energy conservation and recycling behavior (ECRB), nutrition management behavior (Nutrition), health environmental sensitivity (HESens), and physical fitness (Fitness). Personal health responsibility (PHResp) is the only variable in which lower and negative values represent individuals who consider their health more of a personal responsibility.

Finally, the last two regressors in Table 6.1, "Price" and "Appearance", are variables that represent the importance of these attributes to the consumers when buying fresh produce. These variables take the value of 1 if individuals think they are not important, 2 if they are moderately important and 3 if they think they are very important when buying fruits or vegetables.

6.2 Regression Results for Consumer's Preference Models

Following the structure of the model presented in Equation (72), and using the regressors described in the last section, Equation (73) presents the structure of the model for consumer's preferences on organic vegetables and fruits.

$$(73)\quad \left.\begin{array}{c}\text{Prob(Prefer Organic}_{ij}\text{=1)} \\ \scriptstyle j=\text{ fruits, vegetables}\end{array}\right\} = \left\{\begin{array}{l} F(\beta_0 + \beta_1 Primary_i + \beta_2 Female_i + \beta_3 Age_i + \beta_4 Kids_i \\ + \beta_5 Education_i + \beta_6 Income_i + \beta_7 EPB_i + \beta_8 Nutrition_i \\ + \beta_9 ECRB_i + \beta_{10} Fitness_i + \beta_{11} HESens_i + \beta_{12} PHResp_i \\ + \beta_{13} Member_i + \beta_{14} Price_i + \beta_{15} Appearance_i + \varepsilon_i) \end{array}\right.$$

The results and coefficients from the logit regressions are presented in Table 6.2 for vegetables and fruits. All the regression models were solved using a Maximum Likelihood Estimator (MLE)[16] as explained in Chapter 4. Age, environmental purchasing behavior (EPB) and health environmental sensitivity (HESens) were found significant at the 0.05 level by using a z-test in both regressions. Increasing EPB and HESens had positive impact on preference and age negative. Personal health responsibility (PHResp) was found significant at the 0.10

[16] All the models in the section 6.3 were solved using the MLE technique available in the statistical program LIMDEP, v. 7.0 and SPSS 11.5.

level in both regressions. In addition for the vegetables model; nutrition, ECRB, and member were significant at the 0.10 level and price at the 0.05 level. These variables had the same sign in the fruit equation and were of similar size, except for price which had a smaller effect. Being primary shopper, being female, having children, the level of education, income, fitness behavior and appearance did not have a significant effect on the preference choice of either organic vegetables or fruits.

Two measures of fit were calculated for the regression models. The Maddala pseudo-R^2 yielded a measure of 0.38 for fresh vegetables and 0.34 for fresh fruits. The normed measure of Cragg and Uhler yielded 0.55 for vegetables and 0.50 for fruits. The percentage of correct predictions was similar for the models, 80.4% in the case of vegetables and 82.4% in the case of fruits. These measures reveal a good fit for limited dependent variable models. To test for the significance level of all the coefficients, a Wald and a Likelihood Ratio (LR) test were conducted. The Wald statistic was 59.19 (vegetables) and 58.68 (fruits) which is much greater than Chi-Square 1% critical level of 30.6 for 15 degrees of freedom. The LR test also yielded a Chi-Squared statistic that rejects the hypothesis that all the coefficients except the intercept are simultaneously equal to zero.

To better interpret these results the marginal effects on the probability of choosing organic vegetables and fruits are calculated. Since Logit models are nonlinear probability models, there are many ways of presenting the effects of a particular variable on the probability of choosing organic vegetables. Some of the standard techniques are presented in Table 6.3. Calculation of the marginal effects was described in Section 5.4.

Table 6.2: Results for Consumer's Preferences Model

	Fresh Vegetables					Fresh Fruits				
	Coeff.	Std.Err.	z-value	P-value		Coeff.	Std.Err.	z-value	P-value	
Constant	5.605	1.598	3.507	0.000	**	4.687	1.501	3.123	0.002	**
Primary	0.032	0.518	0.061	0.951		-0.157	0.495	-0.318	0.750	
Female	-0.153	0.380	-0.402	0.687		0.165	0.361	0.457	0.648	
Age	-0.050	0.014	-3.710	0.000	**	-0.046	0.013	-3.585	0.000	**
Kids	-0.610	0.440	-1.384	0.166		-0.339	0.423	-0.802	0.423	
Education	0.194	0.216	0.898	0.369		0.151	0.208	0.728	0.466	
Income	-0.083	0.060	-1.376	0.169		-0.069	0.058	-1.182	0.237	
EPB	1.376	0.236	5.825	0.000	**	1.270	0.222	5.732	0.000	**
Nutrition	0.342	0.200	1.709	0.088	*	0.300	0.190	1.580	0.114	
ECRB	-0.315	0.174	-1.814	0.070	*	-0.248	0.166	-1.490	0.136	
Fitness	0.287	0.179	1.604	0.109		0.147	0.173	0.847	0.397	
HESens	0.615	0.183	3.355	0.001	**	0.503	0.170	2.950	0.003	**
PRResp	-0.332	0.193	-1.720	0.085	*	-0.319	0.184	-1.731	0.083	*
Member	0.832	0.492	1.692	0.091	*	0.685	0.464	1.474	0.140	
Price	-0.709	0.334	-2.126	0.033	**	-0.490	0.318	-1.537	0.124	
Appearance	-0.493	0.325	-1.516	0.129		-0.422	0.313	-1.348	0.178	
Measures of fit:	R^2_{ML} =	0.382	$R^2_{C\&U}$ =	0.555		R^2_{ML} =	0.348	$R^2_{C\&U}$ =	0.505	
Log likelihood function:				-102.1368		-109.482				
Restricted log likelihood:				-168.889		-168.3761				
Chi-squared[a] [Critical χ^2 (0.01, 15)= 30.578]:				133.504		117.788				
Wald Statistic[a] [Critical χ^2 (0.01, 15)= 30.578]:				59.192		58.683				

Fresh Vegetables — **Frequencies of actual & predicted outcomes[b]**

Actual\Predicted	0	1	Total
0	66	30	96
1	20	139	159
Total	86	169	255(80.4%)

Fresh Fruits — **Frequencies of actual & predicted outcomes[b]**

Actual\Predicted	0	1	Total
0	64	31	95
1	14	146	160
Total	78	177	255(82.4%)

*** Significant at the 0.05 level*

** Significant at the 0.10 level*

a: testing for b1=b2=b3=b4=b5=b6=b7=b8=b9=b10=b11=b12=b13=b14=b15=0

b: y^=1 if P> 0.5, the cutoff point for predicted values is 0.5

Table 6.3: Marginal Effects for Consumer's Preferences

	Fresh Vegetables					Fresh Fruits				
	(A)	(B)	(C)	(D)	(E)	(A)	(B)	(C)	(D)	(E)
Variable	*Marginal at the means*	*Marginal on average* [a]	*Discrete 0-->1*	*Centered* Δ1	*Base Values* [b]	*Marginal at the means*	*Marginal on average* [a]	*Discrete 0-->1*	*Centered* Δ1	*Base Values* [b]
Primary	0.007	0.004	0.007	---	1.00	-0.036	-0.020	-0.035	---	1.00
Female	-0.036	-0.018	-0.035	---	1.00	0.038	0.021	0.039	---	1.00
Age	-0.012	-0.006	---	-0.012	42.00	-0.011	-0.006	---	-0.010	42.00
Kids	-0.142	-0.073	-0.149	---	0.00	-0.078	-0.044	-0.081	---	0.00
Education	0.045	0.023	---	0.046	3.00	0.035	0.020	---	0.035	3.00
Income	-0.019	-0.010	---	-0.019	5.50	-0.016	-0.009	---	-0.016	5.50
EPB	0.321	0.165	---	0.311	0.00	0.291	0.164	---	0.284	0.00
Nutrition	0.080	0.041	---	0.080	0.00	0.069	0.039	---	0.069	0.00
ECRB	-0.074	-0.038	---	-0.073	0.00	-0.057	-0.032	---	-0.057	0.00
Fitness	0.067	0.034	---	0.067	0.00	0.034	0.019	---	0.034	0.00
HESens	0.143	0.074	---	0.142	0.00	0.115	0.065	---	0.115	0.00
PRResp	-0.077	-0.040	---	-0.077	0.00	-0.073	-0.041	---	-0.073	0.00
Member	0.194	0.100	0.167	---	0.00	0.157	0.088	0.138	---	0.00
Price	-0.165	-0.085	---	-0.160	2.50	-0.112	-0.063	---	-0.112	2.50
Appearance	-0.115	-0.059	---	-0.115	2.50	-0.097	-0.055	---	-0.096	2.50

a: Base values- calculated at the values for each individual
b: Base values- continuous variables held at means, discrete held at mode

Given the nonlinearity of logit models it is necessary to choose base values which allow comparison of the marginal effects at specific levels for the variables. The base values in this study correspond to the means in the case of continuous variables and the mode in the case of discrete. Then, the base case (typical individual) for comparing the results is a female consumer who is the primary shopper age 42, no children, a 4 year college degree, an average household income of $55,000 and not a member of any environmental organization. Note that the values for the components from factor analysis are held at their means, while the attribute variables (Price and appearance) are held at their modes.

The (A) columns present the marginal effects calculated at the means/mode by calculating the partial derivative of the probability function for a particular variable holding the rest of the variables at base values. The (B) columns present the marginal effects calculated at the average; that is, calculating the probability for each observation and taking the average over all of them. The marginals on the average are almost half of the marginals at the mean/mode as result of the nonlinearity of the logit model. For this reason calculating the centered change around the means (D Columns) for continuous variables and the 0 to 1 change (C Columns) for discrete, is helpful. The partial change in probability of choosing organic vegetables and fruits decreases with age. The marginal effect of age is equal to -0.012 for vegetables and -0.010 for fruits which can be interpreted as a reduction of 1 % for each additional year. That is a 52 year old individual (10 years over the base case of 42) has a 10% lower probability of preferring organic vegetables and fruits.

The marginal effects for environmental purchasing behavior are similar in the models. The marginal for vegetables is 0.31 and 0.28 for fruits. These marginal effects can be interpreted as the partial changes in the probability as a result of a small change in the score for EPB. Using the discrete centered change, a change of one unit (one standard deviation) around the mean[17] raises the probability of choosing organic vegetables by 31% (28% for fruits). Consumers with the highest environmental purchasing behavior, who demand products that cause less environmental damage, have the highest probability of preferring organic vegetables and fruits.

Health environment sensitivity (HESens) also has an important impact on preferences for organic produce with a marginal effect of 0.14 and 0.12 for vegetables and fruits respectively. That is, a unit increase in the factor scores for consumers who care more about their water quality and are worried about the presence of chemicals in their food increases the probability of preferring organic vegetables about 14% and fruits 12%.

Personal health responsibility (PHResp) also has a significant effect on the preferences for organic produce. People who have a highest level of personal concern towards their health (lower scores for PHResp) have higher probability of preferring organic produce. The partial change in the probability for a change in PHResp yields a 8% increase in the chances of people preferring organic vegetables and 7% in fruits.

Another variable with an important effect on preferences for organic vegetables is the importance of price with a marginal effect of -0.16. A change from a person who cares moderately about price to one which rates this attribute very

---

[17] Recall that EPB has mean 0, and ranges from -2.97 to 1.79.

important lowers the probability of preferring organic vegetables by 16%[18]. The marginal effect for fruits is only about 1/3 less, -0.11 but was not found significant.

Three other variables were found significant at the 0.10 level for organic vegetables but no fruits, though the coefficients were of nearly the same size in the fruit preference model. Nutrition was found significant with a marginal effect centered at the means of 0.08. One standard deviation increase in the factor scores for consumer who care more about their nutrition (having a daily balanced diet, eating fruits and vegetables and avoiding high levels of cholesterol) raises the probability of preferring organic vegetables 8%. Belonging to an environmental organization raises the probability of preferring organic vegetables 17%. Finally, energy conservation and recycling behavior (ECRB) was found to have a marginal effect of -0.07 indicating a negative relationship with the probability of choosing organic vegetables.

Two other variables that have fairly large marginal effects, but were not found significant, are education and children in the household. Having the highest level of education might increase the probabilities of choosing organic about 4% for fruits and 5% for vegetables. Having children reduces the probability 15% and 8% respectively. The rest of the variables, primary shopper, female, income, fitness and appearance were not significant and their marginal effects were found to be lower than 0.08.

[18] The mean of PRICE is 2.4 which means that most consumers rate price as moderately or very important.

6.3 Regression Results for Consumer's Purchasing Models

To determine what distinguish consumers who actually purchase from those who merely prefer organic produce, this section analyzes the results from the regressions including only those individuals who indicated preference for organic. Controlling for preferences (i.e. working only with people who preferred organic), allows the researchers to determine which variables increase organic purchases. For this reason, the sample set is smaller than the one used in the preference models analyzed in the last section[19]. The regression results for the logit models in which the dependent variable is the proportion of organic vegetables and fruit purchases by the individual are presented in Table 6.4. As seen in Equation (74), the structure of the model is the same as for preferences. The same set of regressors, described in Table 6.1, are used.

The two dependent variables take the value of 1 if more than 50% of the individual's vegetable (fruit) purchases are organic and 0 otherwise. The sample proportion of respondents with vegetable purchases greater than 50% organic is 70% and for fruits is 68%[20]. It can be expected that individuals who reported purchasing organic vegetables and fruits in a proportion greater than 50% are more committed to organic products.

[19] The sample set for organic preference models included 255 observations. For purchase models the sample included 158 observations for vegetables and 159 for fruits.
[20] These proportions correspond to the individuals included in the analysis (individuals who preferred organic fruits and vegetables). The proportion corresponding to the whole sample is 43% and 42% for vegetables and fruit purchases respectively.

$$(74)\quad \underset{j=\text{ fruits, vegetables}}{\text{Prob(Purchase Organic}_{ij}=1)} = \left\{ \begin{array}{l} F(\beta_0 + \beta_1 \Pr imary_i + \beta_2 Female_i + \beta_3 Age_i + \beta_4 Kids_i \\ + \beta_5 Education_i + \beta_6 Income_i + \beta_7 EPB_i + \beta_8 Nutrition_i \\ + \beta_9 ECRB_i + \beta_{10} Fitness_i + \beta_{11} HESens_i + \beta_{12} PH\operatorname{Re}sp_i \\ + \beta_{13} Member_i + \beta_{14} \Pr ice_i + \beta_{15} Appearance_i + \varepsilon_i) \end{array} \right.$$

The coefficients and statistical results are presented in Table 6.4. As in the case of the preference models, age, environmental purchasing behavior and health environment sensitivity were found significant for both the vegetable and fruits choice. Female, income and nutrition were significant only in the case of fruits and PHresp was significant only in the case of vegetables. The price variable is now significant at 0.05 level for both models, while income is significant for vegetables at 0.10 level and fruits at 0.05. The rest of the variables primary, kids, education, ECRB, fitness, member and appearance have P-values greater than 0.10.

The purchase models presented smaller fits in the regressions than the preference models as result of the smaller sample set. The vegetable purchase model has a Maddala R^2 of 0.324 (0.315 for fruit purchases) and a percentage of correct predictions of 83% and 82% respectively, with a higher percentage correct for vegetable purchases than was found for preferences. Cragg and Uhler measures presented fits of 0.496 and 0.471 for vegetables and fruits respectively. As before the null hypothesis that all the regressors except the intercept are equal to zero is rejected in both models. The marginal effects were calculated in the same way as in the preference models. Marginal effects at the means, on average, discrete changes and centered changes were calculated as seen on Table 6.5.

Table 6.4: Results for Consumer's Purchasing Model

	Fresh Vegetables					Fresh Fruits				
	Coeff.	**Std.Err.**	**z-value**	**P-value**		**Coeff.**	**Std.Err.**	**z-value**	**P-value**	
Constant	7.265	2.252	3.227	0.001	**	5.610	2.021	2.776	0.005	**
Primary	-0.995	0.776	-1.283	0.200		-0.653	0.714	-0.915	0.360	
Female	-0.680	0.569	-1.196	0.232		-0.955	0.538	-1.775	0.076	*
Age	-0.032	0.019	-1.657	0.098	*	-0.033	0.018	-1.889	0.059	*
Kids	0.214	0.653	0.327	0.744		-0.361	0.553	-0.653	0.514	
Education	-0.254	0.297	-0.855	0.392		-0.013	0.278	-0.047	0.962	
Income	-0.136	0.077	-1.757	0.079	*	-0.152	0.075	-2.018	0.044	**
EPB	1.971	0.429	4.593	0.000	**	1.853	0.402	4.608	0.000	**
Nutrition	0.424	0.313	1.353	0.176		0.722	0.291	2.480	0.013	**
ECRB	0.062	0.280	0.223	0.824		0.123	0.266	0.464	0.643	
Fitness	-0.007	0.267	-0.027	0.979		-0.025	0.247	-0.102	0.919	
HESens	0.690	0.291	2.375	0.018	**	0.448	0.256	1.747	0.081	*
PRResp	-0.538	0.266	-2.023	0.043	**	-0.098	0.259	-0.378	0.706	
Member	0.142	0.555	0.256	0.798		0.643	0.565	1.138	0.255	
Price	-1.204	0.464	-2.594	0.009	**	-1.065	0.431	-2.471	0.013	**
Appearance	-0.072	0.436	-0.166	0.868		0.126	0.412	0.306	0.759	
Measures of fit:	R^2_{ML} = 0.324		$R^2_{C\&U}$ = 0.496			R^2_{ML} = 0.315		$R^2_{C\&U}$ = 0.471		
Log likelihood function:				-60.12756		-65.86672				
Restricted log likelihood:				-93.45905		-98.20373				
Chi-squared[a] [Critical χ^2 (0.01, 15)= 30.578]:				66.663		64.674				
Wald Statistic[a] [Critical χ^2 (0.01, 15)= 30.578]:				32.592		32.857				

Frequencies of actual & predicted outcomes[b] (Fresh Vegetables)

Actual\Predicted	0	1	Total
0	27	17	44
1	10	104	114
Total	37	121	158(82.9%)

Frequencies of actual & predicted outcomes[b] (Fresh Fruits)

Actual\Predicted	0	1	Total
0	32	17	49
1	12	98	110
Total	44	115	159(81.8%)

*** Significant at the 0.05 level*

** Significant at the 0.10 level*

a: testing for b1=b2=b3=b4=b5=b6=b7=b8=b9=b10=b11=b12=b13=b14=b15=0

b: y^=1 if P> 0.5, the cutoff point for predicted values is 0.5

Table 6.5: Marginal Effects for Consumer's Purchasing Models

	Fresh Vegetables					**Fresh Fruits**				
	(A)	**(B)**	**(C)**	**(D)**	**(E)**	**(A)**	**(B)**	**(C)**	**(D)**	**(E)**
Variable	*Marginal at the means*	*Marginal on average* [a]	*Discrete 0-->1*	*Centered* $\Delta 1$	*Base Values* [b]	*Marginal at the means*	*Marginal on average* [a]	*Discrete 0-->1*	*Centered* $\Delta 1$	*Base Values* [b]
Primary	-0.239	-0.114	-0.243	---	1.00	-0.151	-0.082	-0.160	---	1.00
Female	-0.164	-0.078	-0.168	---	1.00	-0.221	-0.120	-0.234	---	1.00
Age	-0.008	-0.004	---	-0.008	42.00	-0.008	-0.004	---	-0.008	42.00
Kids	0.051	0.024	0.052	---	0.00	-0.083	-0.045	-0.079	---	0.00
Education	-0.061	-0.029	---	-0.062	3.00	-0.003	-0.002	---	-0.003	3.00
Income	-0.033	-0.016	---	-0.033	5.50	-0.035	-0.019	---	-0.035	5.50
EPB	0.474	0.225	---	0.443	0.00	0.428	0.233	---	0.406	0.00
Nutrition	0.102	0.048	---	0.102	0.00	0.167	0.091	---	0.165	0.00
ECRB	0.015	0.007	---	0.015	0.00	0.028	0.016	---	0.028	0.00
Fitness	-0.002	-0.001	---	-0.002	0.00	-0.006	-0.003	---	-0.006	0.00
HESens	0.166	0.079	---	0.165	0.00	0.103	0.056	---	0.103	0.00
PRResp	-0.129	-0.061	---	-0.129	0.00	-0.023	-0.012	---	-0.023	0.00
Member	0.034	0.016	0.035	---	0.00	0.149	0.081	0.157	---	0.00
Price	-0.290	-0.138	---	-0.288	2.50	-0.246	-0.134	---	-0.249	2.50
Appearance	-0.017	-0.008	---	-0.017	2.50	0.029	0.016	---	0.029	2.50

a: Base values- calculated at the values for each individual
b: Base values- continuous variables held at means, discrete held at mode

As before age was significant with a marginal effect of -0.008 for both vegetables and fruits, indicating that for each additional 10 years the probability of being a true organic shopper decreases by 8% vs. 10% as in the preference models. These results suggest that younger consumers are more likely to prefer and to purchase organic fresh produce, but that given preference age is somewhat less important.

As in the case of the preference model, consumers with the highest environmental purchasing behavior have the highest probability of buying organic produce. A one unit increase in the factor score for EPB, increases the probability by 44% and 41% for vegetables and fruits respectively. The effect of EPB is substantially higher in the case of purchases than preferences for organic vegetables (31%) and fruits (28%).

From the set of health related factors, only health environment sensitivity (HESens) was found statistically significant for both models. The centered marginal effects in the case of vegetables and fruits were 0.165 and 0.103 respectively for each additional unit observed around the mean in the factor scores. Individuals that care more about their water quality and are worried about the presence of chemicals in the food have a higher probability of purchasing organic vegetables and fruits.

Contrary to the preference models, price was found highly significant in explaining organic purchases for both vegetables and fruits. The marginal effects for price are -0.29 and -0.25 in the case of vegetables and fruits respectively. In other words, the change from moderately to very important in the perception of price can lead a consumer to be 29% less likely to purchase organic vegetables and 25% less likely for fruits. In the last section, price was found significant only for vegetables preferences with a marginal effect of 16%. This is an expected result, since expressing preferences does not automatically means that the individual will purchase the product. It is when consumers are actually making purchasing decisions that price is fully taken into account, as the higher marginal effects appear to confirm.

Household income, not significant in preferences, was significant in purchases of organic produce. The centered marginal effect for fruits was found to be -0.035 and for vegetables -0.033. For each additional $10,000 of household income, the probability of purchasing organic vegetables and fruits declines by about 4%. It is important to highlight that income is a significant variable negatively affecting the probability of purchasing organic produce but not preferences.

Personal health responsibility (PHResp) was significant for vegetable purchases, and nutrition only for fruit purchases. The centered marginal effect is -0.13 for each additional unit observed around the mean in the factor scores. Individuals that have a highest level of personal responsibility towards their health are more likely to buy organic vegetables (lower values of PHResp correspond to higher probability of purchasing organic vegetables). The nutrition variable measures individuals concern about having a nutritionally balanced diet, who avoid high cholesterol diets, and who are interested in information about their health. A standard deviation increase (one unit increase) raises probability for purchasing organic fruits 17%. These results imply that while vegetable choice was related to personal health responsibility issues, fruit selection is influenced by nutrition concerns for the individuals in the sample. However, these variables have an obvious relationship and will be discussed in the conclusions.

Female was also only significant in influencing purchasing level of fruits. The marginal effect for female holding the rest of the variables at their base values is -0.23. This implies that, given the sample, being a female consumer reduces the probability of buying organic fruits by 23%[21]. This result suggests that females are less likely to buy organic fruits for reasons different then being the primary shopper in the household or being more price conscious since both "primary" and "price" can account for these effects.

Neither primary, having children, member, having higher levels of education, energy conservation and recycling behavior, fitness, or appearance were found to have significant effects on the probability of purchasing organic produce at the 0.10 level.

6.4 Alternative Models Tested

Two other methods for approaching binomial choice models besides the multinomial logit were tested. The binomial probit and the linear probability model (LPM) were tested using the same set of variables used in the logit approach. In the case of the LPM the results obtained were discarded since they suffered the limitations of using this approach, having a linear probability function imposes a

[21] In the case of vegetables the p-value of Female is 0.198. Even tough is not significant the coefficient was also negative.

functional form problem that allows the probabilities not to add to one. In the case of the probit, similar results for the marginal effects for the independent variables were obtained as in the logit. Because the simplicity of solving the probability functions and calculating the marginal effects, the logit model was preferred.

Other demographic variables were also tested, job, number of members in the household, area where the individual grew up, and owning current residence. Other variable forms were also tested. In the case of education four alternative coding schemes were tested: as categorical dummies, treatment as a continuous variable, based on a logarithmic function and implied years of education. In no case was education found significant. Age and income were also examined in logarithmic forms without improvement. The variables presented in the final model were chosen based on statistical and theoretical considerations.

7. Conclusions

The main objective of this research was to identify the main factors explaining consumer's choice of organic versus conventionally grown produce. The principal determinants of organic preference choice and purchase decisions for the sample used in this study are presented in Table 7.1. Socio demographic characteristics as well as environmental and health motives are important in explaining the preference and purchasing choice of organic.

Table 7.1: Marginal Effects on the probability of preferring and purchasing organic

			Preference Model		**Purchase Model**	
			Vegetables (n=255)	**Fruits (n=255)**	**Vegetables (n=158)**	**Fruits (n=159)**
Demographics	Primary					
	Female	Female=1				* (-23%)
	Age	+10 years	** (-12%)	** (-10%)	* (-8%)	*(-8%)
	Kids					
	Education					
	Income	+$10,000			* (-3.3%)	** (-3.5%)
	Member	Member=1	*(17%)			
Environmental Factors	Environmental Purchasing Behavior	+ 1	**(31%)	**(28%)	**(44%)	**(41%)
	Energy Conservation and Recycling Behavior	+ 1	* (-7%)			
Health Factors	Nutrition Management Behavior	+1	*(8%)			**(17%)
	Physical Fitness					
	Health Environmental Sensitivity	+1	**(14%)	**(11%)	**(17%)	*(10%)
	Personal Health Responsability	-1	*(8%)	*(7%)	**(13%)	
Attributes	Price	+1	**(-16%)		**(-29%)	**(-25%)
	Appearance					

a: Marginal effects: centered around the mean for continuous and discrete change for dummies
*** Coefficients significant at the 0.05 level*
** Coefficients significant at the 0.10 level*

Consumer preference models are affected primarily by age, environmental purchasing behavior (EPB), health environmental sensitivity (HESens) and personal health responsibility (PHResp). Younger consumer's and individuals exhibiting concern about the effect of their purchases on the environment are the most likely to prefer organic. Moreover, individuals concern about the effect of chemicals in food and water on their personal health, and who reported their health was their responsibility, are also more likely to prefer organic. Environmental motives were found to have a greater influence on preferences for organic produce than health motives. As seen in Table 7.1, environmental purchasing behavior has a marginal affect on the probability of preferring organic vegetables and fruits of 30% and 28% respectively. Health factors represented by health HESens and PHResp exhibit a lower impact on the likelihood of preferring organic over the variation observed.

Purchasing likelihood is also driven primarily by environmental motives. The marginal effects for EPB in the case of purchases for vegetables and fruits are 45% and 42% respectively. Environmental purchasing behavior was found to have a larger impact on purchases than on preferences as shown by the even larger marginal effects. Health related factors, health environmental sensitivity, personal health responsibility (Vegetables) and nutrition (Fruits) are also important in influencing consumer's purchases for organic produce in a proportion greater than 50%, but given the results found for this sample, environmental motives are more important. Income and consumer's sensitivity to price were also significant in explaining consumer's purchases for both fruits and vegetables. The more important price is to an individual the less likely they are to purchase mostly organic produce. On the other hand, individuals from higher income groups have a lower probability of purchasing 50% or more organic produce. While price sensitivity is a clear attribute due to the higher prices for organic produce, the income result is likely to be related to other demographic aspects.

The importance of environmental purchasing behavior in both models, particularly in explaining purchases, indicates the importance that consumer's behavior towards the environment has on explaining organic choice. Loureiro, McCluskey and Mittelhammer (2001) found in their analysis of the choice of organic and eco labeled apples in Portland, OR that their variable used to represent environmental[22] concerns was significant with a marginal effect of 0.36 and higher than food safety concerns with 0.16. Blend and van Ravenswaay (1999), also found in their analysis of eco-labeled apples that their variable representing environmental[23] concern was significant with a marginal effect of 0.54 though they did not find health significant. The results of this study confirm that concern for the environment is a very important motive in determining both preferences and purchases. This study finds that environmental motives behind organic choice are in step with their behavior towards environmentally responsible products (products that cause less pollution, that do not harm the environment, dolphin-safe tuna, etc) rather than with energy conservation and recycling.

Younger people were found to be more likely to prefer and purchase organic fresh produce. The effect of age was found to be greater in influencing preferences than purchases. The negative relationship between age and organic preferences

[22] Trade off variable measuring the importance of environmental protection versus jobs.
[23] If consumers think that by purchasing eco-labeled apples they can protect the environment.

corroborates the Govindasamy and Italia (1999) results, but Thompson and Kidwell (1998) did not find age to be a significant variable, and Thompson's (1998) survey of other studies on organic choice reported mixed results. Young consumers might see use of organic produce as a way to improve the environment in ways beyond those explained by the environmental variables included in the model. They might perceive greater benefits or have a higher level of social consciousness. They may also be more optimistic about the impact their behavior will have.

The most important health factor influencing preference and purchases of organic produce was health environmental sensitivity. Given the sample, respondents that reported being concern about the drinking water quality and who worry that there are harmful chemicals in food are more likely to prefer and make the majority of their produce purchases organic. Consumers have repeatedly reported reduced use of pesticides, or use of no pesticides, as one of the main reasons to purchase organic products as discussed in Chapter 2. Huang (1996) found that concern about pesticide residues is very important in determining preference for organic produce.

Consumers that are more price oriented are less likely to prefer vegetables and to purchase organic produce (both vegetables and fruits). The effect of price was found to be higher in influencing purchases than preferences for vegetables. Purchase decisions are naturally more likely to be related to price considerations. The negative relationship between price and organic preference was also found by Huang (1996) and Thompson and Kidwell (1998) in organic purchase likelihood. Blend and van Ravenswaay (1999) found price conscious consumers to be less likely to purchase eco-labeled apples. Wang and Sun (2003) found in their conjoint analysis that price was the most important reason stated by consumers for why they have never purchased organic food. All studies that have included a price variable have reached the same conclusion, consumers who are concerned with higher prices or premiums are less likely to prefer and purchase organic produce. Huang in the acceptance of organic produce, Thompson and Kidwell in the choice of organic fresh produce, Blend and van Ravenswaay in the demand for ecolabeled apples and Wang and Sun in their analysis of organic apples and milk.

Despite the price premium that organic produce has over conventional, lower income groups are more likely to buy organic produce in a proportion greater than 50%. This study found that the coefficient for income was negative and significant in purchasing produce but not significant in affecting preferences. Literature on the topic has found conflicting results about the effect of income on preferences and purchases. Studies on organic preferences and willingness to pay price premiums found income being significant and positive in affecting consumer's preferences. Loureiro and Hine (2001) found upperclass consumers (high income and high education levels) more likely to pay a premium for organic potatoes. A positive relationship was also found by Govindasamy and Italia (1999) on their study on willingness to pay a premium for organic produce and income. Perhaps lifestyles or career choices related to the principles of organic farming have led to lower household incomes. It is important to note that the studies that found income having positive effect on organic choice used preference and willingness to pay questions. This study examined the actual percentage or organic purchases reported by the household rather than looking at a single item. Given this contrast, it may be inferred that lower income groups, have increased likelihood of organic purchase, but with no significant difference among those who prefer. However, the result in this study may be sample related.

There were four variables in the preference models that particularly influence choice for organic vegetables: member, ECRB, nutrition and price. However, the coefficients for these variables, except for price, are similar to the ones found in the fruits model. Being member in an environmental organization raises the probability of preferring organic but it was not significant in the purchase models. Members of environmental organizations may be more knowledgeable about organic products and benefits which may increase preferences for organic. Price sensitivity and energy conservation and recycling behavior (ECRB) decrease the likelihood for preferring organic vegetables. ECRB denotes an efficient use of resources by recycling paper, cans and bottles and reducing energy consumption (saving electricity, buying efficient light bulbs and using rechargeable batteries). This result was not expected since it implies that consumers who are concern more about recycling and energy conservation are less likely to prefer organic vegetables. Perhaps consumers who seek efficiency in the use and disposal of materials and energy, do not perceive organic farming as an efficient process. It is possible that these consumers might have the impression that organic farming, by reducing the use of pesticides and fertilizers, is not as efficient as conventional farming. Certainly this phenomenon needs more attention to be fully understood.

Personal health responsibility (PHResp) for both fruits and vegetables was found to be important at a 10% level of significance in preference models. In the case of purchase models PHResp was significant for fruits only and nutrition significant for vegetables only, both at a 5% level. Perhaps consumers taking greater responsibility for their health are motivated to consume vegetables and consumers with greater nutritional concerns are motivated to consume fruits. This tradeoff may have something to do with their similar implication regarding personal health responsibility. It may also distinguish general perception about fruits and nutrition or perhaps an important presence of vegetarians in the sample. For vegetarians the importance of vegetables in their diets might well be perceived as an overall health responsibility factor.

Being female was found significant only in decreasing the probability of purchasing 50% or more of organic fruits. Perhaps, females are less likely to purchase organic fruits because of other reasons not directly accounted for in the model. It is possible that females, while considering fruits important, are less likely to restrict themselves to organic sources because conventional fruits are more convenient and readily available. Gender was not found significant in affecting preferences for either product. Huang (1996), Thompson and Kidwell (1998) and Loureiro and Hine (2001) also found that gender was not an important variable in explaining choice of organic produce. On the other hand Govindasamy and Italia (1999), and Loureiro, McCluskey and Mittelhammer (2001) found evidence of gender being an important variable in explaining preferences for organic food with female having a positive effect. Thompson (1998) in his review of the principal studies conducted before 1998, concluded that there is limited evidence suggesting that gender per se contributes to explaining differences in organic purchase behavior.

As mentioned in earlier chapters, a question regarding the importance of buying Free of GMO products was included. This variable was found to be correlated with both health and environmental variables. Analysis showed partial correlation coefficients of 0.48 between free of GMO and the health environment sensitivity factor and 0.56 between free of GMO and environmental purchasing behavior. When free of GMO is included as an individual variable in the preference and purchase models, HESens becomes insignificant, lowering the marginal effects of both EPB and HESens while results for the rest of the variables are similar[24] (See Appendix 7). This result suggests that free of GMO was capturing some of the explanatory power of the health and environmental variables in explaining both preference and purchase likelihood[25]. By including the Free of GMO variable in the factor analysis it was found to be correlated in a higher degree with Environmental Purchasing Behavior (EPB) rather than with HESens. Factor analysis proved to be useful by classifying this variable in the environmental factor (EPB) and reducing the collinearity between free of GMO and the rest of the variables.

As discussed in the last chapter, primary household shopper, children, educational level, fitness and appearance were insignificant in all of the discrete choice models. In the case of presence of children in the household, research studies also suggests mixed results. This study, found that children are not significant in explaining the choice of organic produce, like Govindasamy and Italia (1999). Being

---

[24] As seen in Appendix 7, when free of GMO is included HESens is not significant at 0.05 level with a marginal effect of 8.3%.

[25] A simple linear regression: Free of GMO = f(HESens, EPB) yielded coefficients of .237 (HESens) and .354 (EPB) with R^2=.38. This also shows that free of GMO is related to, but not entirely explained by, environmental and health concerns.

the primary shopper of the household was also insignificant in both the preference and purchase models. Govindasamy and Italia's study also found primary shopper insignificant in willingness to pay for organic.

The literature reveals mixed evidence about the effect of the level of education on the choice of organic food. Thompson and Kidwell (1998) found a negative relation between having an advanced college degree and the choice of organic fresh produce. Meanwhile, Govindasamy and Italia (1999) found a negative relationship between education level and willingness to pay premium for organically grown produce. On the other hand Loureiro and Hine (2001) found a positive relation between education and demand for organic potatoes. Thompson (1998) suggests the existence of evidence of a positive relationship between these two variables. The results found in this study are not statistically significant for education but are positive in both preferences and purchase likelihood. Given the mixed results found in the literature, and the results obtained in this analysis, no definite conclusions can be drawn about the effect of education on the preference and purchasing choice of organic produce.

The statistical significance of the variables in Table 7.1 is partially explained by the cross tabulations presented in Section 5.1. Overall only minor demographic differences were noted between organic fruit and vegetable buyers. Given the sample, individuals who prefer organic vegetables also prefer organic fruits explaining the similar demographics and econometric results. Neither primary shopper, level of education, or number of children in the household showed different patterns in the frequency analysis by choice of organic produce. This was reflected in the statistical significance of these variables in the regressions. On the other hand, age and income, showed distinctive patterns, explaining why these variables were found statistically significant.

The factor analysis identified two dimensions in environmental concerns (environmental purchasing behavior and energy conservation and recycling behavior) and four dimensions in health concerns (nutrition, health environment sensitivity, physical fitness and personal health responsibility). The factor scores that resulted from these dimensions, particularly EPB and HESens, proved to be very important in the regression models for fresh produce. The use of behavioral questions and the use of factor analysis to extract the principal components were important for the purposes of this research and in improving model efficiency.

Sellers of organic produce in the venues analyzed by this study should recognize the importance that environmental motives have on the choice and purchase of fresh vegetables and fruits. Marketing strategies should stress the environmental factor embedded in organic products as a key selling point. In contrast to health claims, national organic standards do not prohibit the use of environmentally friendly claims. Moreover, organic fresh marketers of the three store venues analyzed should distinguish their product from eco-labeled products by stressing the fact that organic seeks the overall harmony of the ecosystem that surrounds the production process and the fact that they do not use GMO products. It is also important to recognize in targeting markets the fact that younger shoppers have a higher probability of purchasing organic produce.

The main limitation of this study is the limited applicability of the results found from this sample to the general population. The sample used in this study deliberately over sampled organic consumers and took place in one geographical region. Broadening the sample size and geographical scope of this study would be useful in testing whether the results found hold in a more general population.

One problem encountered during the development of the survey was the difficulty of designing questions that would not require open ended responses to ease the use of Tablet PC's. The alternative chosen for ranking the most important attributes when buying fresh produce confused respondents. This study deliberately omitted the inclusion of open ended questions regarding the principal factor of why consumers demand organic produce. This information would have been interesting to compare with the regression results.

There are three avenues suggested for future research. First, the relationship between organic preferences and/or purchases and genetically modified organisms should be studied more deeply. Survey questions need to be specifically modeled to determine whether people consume organic products because GMO products might have in impact on their health or on the environment or whether GMO has shifted some to organic buying. Similar studies should be undertaken to study the principal factors affecting consumer's choice of other products in addition to fresh produce. Soft drinks, snacks, meat/poultry and many others offer a wide set of choices. Finally, the model should be applied to other geographical regions to test if the same motivations are found affecting consumer's choice of fresh produce.

Bibliography

Baker, G., and P. J. Crosbie, "Measuring Food Safety Preferences: Identifying Consumer Segments", Journal of Agriculture and Resource Economic, No.18, 277-287 (1993).

Ben-Akiva, M., and S. Lerman., "Discrete Choice Analysis: Theory and Application to Travel Demand", Cambridge, The MIT Press (1985).

Blend, J., and E. van Ravenswaay, "Measuring Consumer Demand for Ecolabeled apples", American Journal of Agricultural Economics, Vol. 81, No. 5, 1072-1077 (1999).

Borooah, V. K., "Logit and Probit: Ordered and Multinomial Models", Sage University Paper No 138, Sage Publications (2002).

Byrne, P., U. Toensmeyer, C. German, and H. Reed Muller, "Analysis of Consumer Attitudes Toward Organic Produce and Purchase Likelihood", Journal of Food Distribution Research, Vol. 22, No. 2, 49-62 (1991).

Byrne, P., R. Bacon and U. Toensmeyer, "Pesticide Residue Concerns And Shopping Location Likelihood", Agribusiness, Vol. 10, No.6, 491-501 (1994).

Dimitri C., and S. Richman, "Organic Market in Transitions", Henry A. Wallace Center for Agricultural & Environmental Policy, Winrock International Policy Studies Report No. 14 (2000).

Dimitri, C., and C. Green, "Recent Growth in the U.S. Organic Food Market", U.S Department of Agriculture, Economic Research Service, Market and Trade Economics Division and Resource Economics Division, Agriculture Information, Bulletin Number 777 (2002).

Eom, Y., "Pesticide Residue Risk and Food Safety Valuation: A random utility approach", American Journal of Agricultural Economics, Vol.76, No. 4, 760-771 (1994).

Estes, E., and V. K. Smith, "Price, Quality, and Pesticide-Related Health Risk Considerations in Fruit and Vegetable Purchases: An Hedonic Analysis of Tucson, Arizona Supermarkets," Journal of Food Distribution Research, Vol. 27, No. 3, 8-17 (1996).

Food Marketing Institute and Prevention Magazine, "Shopping for Health 1997: Balancing Convenience, Nutrition and Taste", Prepared by Princeton Survey Research Associates (1997).

Glaser, L., and G. Thompson, "Demand for Organic and Conventional Frozen Vegetables", Selected Paper presented at Annual Meeting of the American Agricultural Economics Association, Nashville, TN, August (1999).

Govindasamy, R., and J. Italia, "Predicting Willingness to Pay a Premium for Organically Grown Fresh Produce", Journal of Food Distribution Research, Vol. 30, No. 2, 44-53 (1999).

Govindasamy, R., J. Italia, and A. Adelaja, "Predicting Willingness-to-Pay a Premium for Integrated Pest Management Produce: A Logistic Approach", Agricultural and Resource Economics Review, Vol. 30, No. 2, 151-159 (2001).

Greene, W.H., "Econometric Analysis", New Jersey, Prentice Hall (2000).

Hartman Group, "Food and the Environment: A Consumer's Perspective, Phase I", Bellevue WA, The Hartman Group (1996).

Holland D., and C. Wessells, "Predicting Consumer Preferences for Fresh Salmon: The Influence of Safety Inspection and Production Method Attributes", Agricultural and Resource Economics Review, Vol. 27, No. 1, 1-14 (1998).

Huang, C. L., "Consumer Preferences and Attitudes towards Organically Grown Produce", European Review of Agricultural Economics, No. 23, 331-342 (1996).

Kline, P., "An Easy Guide to Factor Analysis", Routledge (1994).

Kraft, F. B., and P.W. Goodell., "Identifying the Health Conscious Consumer", Journal of Health Care Marketing, Vol. 13, No. 3, 18-25 (1993).

Long, J. S., "Regression Models for Categorical and Limited Dependent Variables", Sage Publications (1997).

Loureiro, M., and S. Hine, "Discovering Niche Markets: A Comparison of Willingness to Pay for a Local (Colorado Grown), Organic, and GMO-Free Product", Selected Paper presented at Annual Meeting of the American Agricultural Economics Association, Chicago, IL (2001).

Loureiro, M. L., J. J. McCluskey, and R. C. Mittelhammer, "Assessing Consumer Preferences for Organic, Eco-labeled, and Regular Apples", Journal of Agricultural and Resource Economics, Vol. 26, No. 2, 404-416 (2001).

Maddala G.S., and C.R. Rao, "Handbook of Statistics 14", North-Holland (1996).

Manski C. F., and D. McFadden, "Structural Analysis of Discrete Data with Econometric Applications", Cambridge, The MIT Press (1981).

Mas-Colell, A., M. Whinston, and J. Green, "Microeconomic Theory", Oxford University Press (1995).

Myers, S., and S. Rorie, "Facts and stats: the year in review. Organics and Natural BusinessNews",http://www.organicandnaturalnews.com/articles/0c1feat1.html/ viewed August 14 (2003).

National Organic Standards Board, http://www.ams.usda.gov/nosb/index.htm, viewed August 14 (2003)

Nicholson, W., "Microeconomic Theory: Basic Principles and Extensions", South-Western/Thompson Learning (2002).

Organic Trade Association, Organic Facts, http://www.ota.com/, viewed August 14, (2003).

Parkwood Research Associates, "Shopping for Organic: Food Shopper's View on Organically Grown Produce, Summary Report, Rodale Press (1994).

Rencher, A.C., "Methods of Multivariate Analysis", Wiley-Interscience (2002).

Roberts, J.A., "Green Consumers in the 1990's: Profile and Implications for Advertising", Journal of Business Research, No. 36 (1996).

Srivastava, M. S., "Methods of Multivariate Statistics", Wiley-Interscience (2002).

Straughan, R. D., and J. A. Roberts, "Environmental segmentation alternatives: a look at green consumer behavior in the new millennium", Journal of Consumer Marketing, Vol. 16, No. 6, 558-575 (1999).

Teisl, M. F., B. Roe, and A. S. Levy, "Eco-Certification:Why It May Not Be a Field of Dreams", American Journal of Agricultural Economics, Vol. 81, No. 5, 1066-1071 (1999).

The Packer, "Fresh Trends 1998", Vance Publishing (1998).

Thompson, G., "Consumer Demand For Organic Produce: What We Know And What We Need To Know", American Journal of Agricultural Economics, Vol. 80, No. 5, 1113-1112 (1998).

Thompson, G., and J. Kidwell, "Explaining the Choice of Organic Produce: Cosmetic Defects, Prices, and Consumer Preferences", American Journal of Agricultural Economics, Vol. 80, No. 2 , 277-287 (1998).

Train, K., "Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand", Cambridge, The MIT Press (1986).

U.S. Census Bureau, "2002 American Community Survey Profile", http://www.census.gov/acs/www/Products/Profiles, viewed November 6, (2003).

United States Department of Agriculture (USDA), "Organic Farming Statistics and Social Science Research", http://www.ers.usda.gov/Briefing/Organic, viewed November 6, (2003).

Varian, R. H., "Microeconomic Analysis", New York, Norton & Company (1992).

Wang, Q., and J. Sun, "Consumer preference and demand for organic food: Evidence from a Vermont survey", Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Totonto-Canada (2003).

Wessells, C. R., J. Johnston, and H. Donath, "Assesing Consumer Preferences for Eco-labeled Seafood: The Influence of Species, Certifier, and Household Attributes", American Journal of Agricultural Economics, Vol. 81, No. 5, 1084-1089 (1999).

Appendices

OREGON STATE UNIVERSITY

Appendix 1: Survey

Agricultural and Resource Economics
On Site Survey

1. How often do you have a serving of fruit (e.g. an apple, a banana)?

- Two or more times a day
- At least once a day
- 4-6 times per week
- 1-3 times per week
- less than once a week

2. Where do you regularly buy fresh fruit and/or vegetables? (Please mark all that apply)

- Supermarkets
- Co-op
- Convenience Stores
- Grocery Store
- Farmers Market
- Directly from farms
- Grown at home

3. How often do you have a serving of vegetables (e.g. a cup)?

- Two or more times a day
- At least once a day
- 4-6 times per week
- 1-3 times per week
- less than once a week

4. How much do you usually spend on <u>groceries</u> per week?

- Less than 10$
- 31-40
- 61-70
- 91-100
- 11$- 20$
- 41-50
- 71-80
- 101-110
- 21$- 30$
- 51-60
- 81-90
- More than 110$

5. How much do you usually spend on fruits per week?

Less than 5$	16-20	31-35	51-55
6$- 10$	21-25	36-40	56-60
11$- 15$	26-30	41-50	More than 60$

6. How much do you usually spend on vegetables per week?

Less than 5$	16-20	31-35	51-55
6$- 10$	21-25	36-40	56-60
11$- 15$	26-30	41-50	More than 60$

7. Please specify what type of products you *usually prefer* when buying the following products (Please mark only one column per item).

	Conventional	Organic	Other Eco-Label	Don't Buy
Coffee / Tea				
Juice				
Dairy Products				
Fresh Vegetables				
Fresh Fruit				
Frozen foods				
Canned foods				
Eggs				
Beer / Wine				
Meat / Poultry				
Snacks / Cereals				
Baby food				

8. Have you purchased products with any of these eco-labels?

- Food Alliance
- Salmon Safe
- Rainforest Alliance
- Marine Stewardship Council
- None that I know of.
- Shade Grown Coffee
- Fair Trade Certified
- Dolphin Safe Tuna
- Other eco-label

9. If the price of conventional apples is 0.99 $/pound, how much would you be willing to pay for a pound of *organic apples*?

$0.79 $0.99 $1.29 $1.59 $1.89 $2.19

10. Are you willing to pay a price premium for *fruit juice* that has been certified as organic?

Yes No

If yes, would you pay a:

20% price premium
40% price premium
60% price premium
80% price premium
more than 40% price premium

11. Are you willing to pay a price premium for *vegetables* that have been certified as organic?

Yes No

If yes, would you pay a:

20% price premium
40% price premium
60% price premium
80% price premium
more than 80% price premium

12. What percentage of your <u>fresh fruit</u> purchases is organic?

0% 1-10% 11-20%21-30% 31-40% 41-50% 51-60% 61-70% 71-80% 81-90% 91-100%

13. What percentage of your <u>vegetables</u> purchases is organic?

0% 1-10% 11-20% 21-30% 31-40% 41-50% 51-60% 61-70% 71-80% 81-90% 91-100%

14. Please rank the reasons you choose not to buy the organic product with 4 as the most frequent reason you choose not to buy, and 1 as the least frequent, give a rank of 0 or skip if not a reason.

a) Organic not readily available	5	4	3	2	1	0
b) Organic priced too high	5	4	3	2	1	0
c) Organic not in good condition	5	4	3	2	1	0
d) Organic not locally grown produce.	5	4	3	2	1	0

15. How important are each one of the following attributes when buying fruits or vegetables - rate your answers from "very important" to "not important" and <u>then rank the top 5 reasons only</u> in the boxes in order of importance with 5 as most important.

	Very Important	Moderately Important	Not Important	Rank				
				5	4	3	2	1
Price								
Product Appearance								
Size								
Taste								
Ripeness								
Locally Grown								
Nutritional Value								
In season								
Pesticide use reduction								
Organically grown								
Free of GMO (Genetically Modified Organism)								

16. How well do the following statements describe your personal behavior on environmental and health issues?

	Always True	Mostly True	Sometimes True	Rarely True	Never True
a) I recycle paper, cans or bottles.					
b) Do you look for the dolphin-safe label when you buy tuna?					
c) I buy "environmentally friendly" products, even if they are more expensive.					
d) I use rechargeable batteries.					

	Always True	Mostly True	Sometimes True	Rarely True	Never True
e) I buy energy efficient light bulbs for my household.					
f) I purchase recycled paper.					
g) I have switched products for environmental reasons.					
h) I have convinced members of my family or friends not to buy some products that are harmful to the environment.					
i) I will not buy a product if the company who sells it is ecologically irresponsible.					
j) I have tried very hard to reduce the amount of electricity I use.					
k) I have purchased products because they cause less pollution.					
l) I try only to buy products than can be recycled					
m) I do not buy household products that harm the environment					
n) I avoid foods containing nitrites or preservatives.					
o) I am interested in information about my health.					
p) My daily diet is nutritionally balanced.					
q) I eat 5 servings of fruits and vegetables a day.					
r) I regularly participate in outdoor activities (walking, biking, etc.).					
s) I try to avoid stressful situations					
t) I try to exercise at least 30 min. a day, three days a week.					

17. Indicate your level of agreement with the following statements.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
a) I read more health-related articles than I did 3 years ago					
b) I worry that there are harmful chemicals in my food.					
c) I'm concerned about my drinking water quality.					
d) Good health takes active participation on my part.					
e) It is the doctor's job to keep me well.					
f) My health is outside my control.					
g) I try to avoid high levels of cholesterol in my diet.					
h) I exercise more than I did three years ago					
i) The health benefits of eating fresh fruits and vegetables far outweigh the health risks from possible pesticide residues.					

18. When buying *fruits and vegetables*, do you feel that the following are reasons for concern? – Rate each component from being "Unconcerned" to being "Very concerned" or "Don't know":

	Unconcerned	Somewhat concerned	Concerned	Very Concerned	Don't Know
Pesticide residues on food	○	○	○	○	○
Preservatives and additives	○	○	○	○	○
Presence of chemicals that cause cancer	○	○	○	○	○
Possible effects of GMO products on your health	○	○	○	○	○

19. When you are purchasing foods, what is the importance of **buying products friendly to the environment** versus **lower costs of food** on a scale of 1 to 10, where **1 means buying products friendly to the environment is all important** and **10 means lowers costs of food is all important**? (CHOOSE ONE)

1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

20. When you are purchasing food, how important is reducing **harmful chemicals** versus **lower cost of food.** Please, place yourself on a scale of 1 to 10, where **1 means reducing harmful chemicals is all important** and **10 means lower food prices are all important**? (CHOOSE ONE)

1	2	3	4	5	6	7	8	9	10
○	○	○	○	○	○	○	○	○	○

21. Have you had any illnesses or food poisoning when consuming the following? (Select any of the following options that apply)

- ○ Fruit Juice
- ○ Oysters or other shellfish
- ○ Red Meat (beef, pork, lamb, etc)
- ○ Salad Dressings (mayonnaise, etc.)
- ○ Dairy Products (cheese, milk, butter, etc.)
- ○ Chicken
- ○ Fish, Shrimp
- ○ Salad Bar
- ○ Fresh Fruit
- ○ Can Food
- ○ Other

THE REST OF THE QUESTIONS ARE YES-NO or multiple choice and cover your shopping habits and demographic information.

22. What is your age group?

- 18 - 24
- 25 - 29
- 30 – 34
- 35 - 39
- 40 - 44
- 45 - 49
- 50 - 54
- 55 - 59
- 60 - 64
- 65 - 69
- More than 70

23. Are you the primary grocery shopper for your household?

- Yes
- No

24. Do you own or rent your current residence? Choose one of the following options:

- Own
- Rent

25. What is the highest level of formal education you have completed? Choose one of the following options:

- High School
- 4 year College Degree
- 2 year College or Technical Degree
- Advanced college degree (i.e.MS, PhD,MD,JD)

26. Was the area where you grew up? Choose one of the following options:

- Urban
- Urban-Rural Mix
- Suburban
- Rural area

27. What range does your total household income fall into? Choose one of the following options:

- $100,000 or more
- $80,000 to $99,999
- $60,000 to $79,999
- $50,000 to $59,999
- $40,000 to $49,999
- $30,000 to $39,999
- $20,000 to $29,999
- Less than $20,000

28. Are you member of an environmental organization?

◯ Yes ◯ No

29. Do you have a job outside the home?

◯ Full Time ◯ Part time

◯ Not employed outside the home

30. Including yourself, how many people in your household are in each of the following age groups?

	1	2	3	4	5	More than 5
Over 65	◯	◯	◯	◯	◯	◯
40 to 64	◯	◯	◯	◯	◯	◯
25 to 39	◯	◯	◯	◯	◯	◯
18 to 24	◯	◯	◯	◯	◯	◯
13 to 17	◯	◯	◯	◯	◯	◯
5 to 13	◯	◯	◯	◯	◯	◯
Less than 5	◯	◯	◯	◯	◯	◯

31. Gender:

◯ Male

◯ Female

That concludes the survey. We appreciate your time.

Have a nice day!

Appendix 2: Table results of principal studies on Organic Preferences

Table 1: Table results of principal studies on Organic Preferences

Variable	Huang (1996)	Thompson and Kidwell (1998)	Blend et al. (1999)	Govindasamy and Italia (1999)	Lourerio et al. (2001)	Govindasamy et al. (2001),	Loureiro and Hine (2001)
Dependent variable	Preference for Organic produce and aceptance of sensory defects	Actual choices of organic fresh produce	Probability of purchase for eco-labeled apples, and quantity purchased	WTP premium for organically grown produce	Choice prefrence among organic, eco-labeled and regular apples	Purchase choice for integrate pest management (IPM)	WTP premium for colorado grown and organic potatoes
Method	Bivariate Probit Model	Bivariate Probit Model	Cragg Double-Hurdle and a Probit Model	Logit	Multinomial logit	Binomial Logit	Multiple bounded probit model
Product	Produce	Fresh Produce	Apples	Produce	Apples		Potatoes
Appearance	X	-					
Income			X	+ (for income groups over 70,000)	X	+ (for income groups over 70,000)	+(upperclass = income + high education)
Age		X	X	+(under 36 years)		+(under 36 years)	-
Gender		X	Female	Female	Female	Female	X
Marital Status							
Education level		- (for Graduate level)	+	-			+(upperclass = income + high education)
Household size				-	-		
KIDS		+		X	+		-
Improve the environment			+ (on quantity)		+		
Food Safety Concern			X(on quantity)		+		
Concern Pesticide Residues	+						
Pesticides Residues on Health				X			
Pesticides Residues on Environ				X		+	
Nutritionally conscious	+						
Price conscious	-	-	-				

WTP= Willingness to pay   *- = negative relationship*
+ = positive relationship   *X= not significant*

Appendix 3: Cross tabulation of demographic variables by venue

Figure 1: Gender by venue

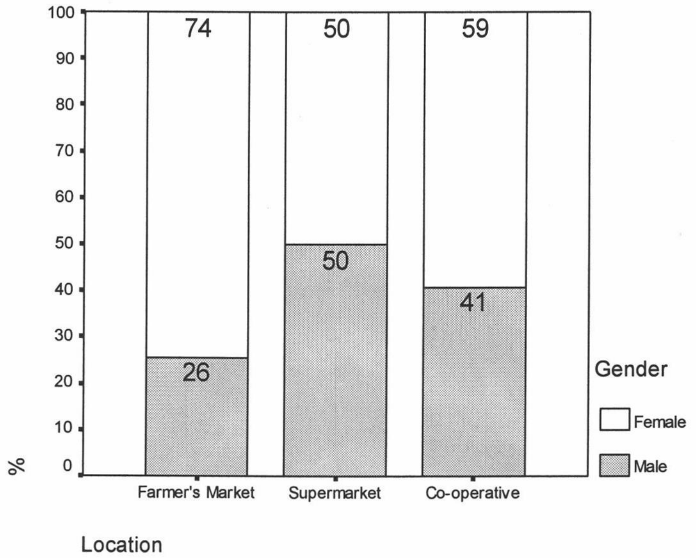
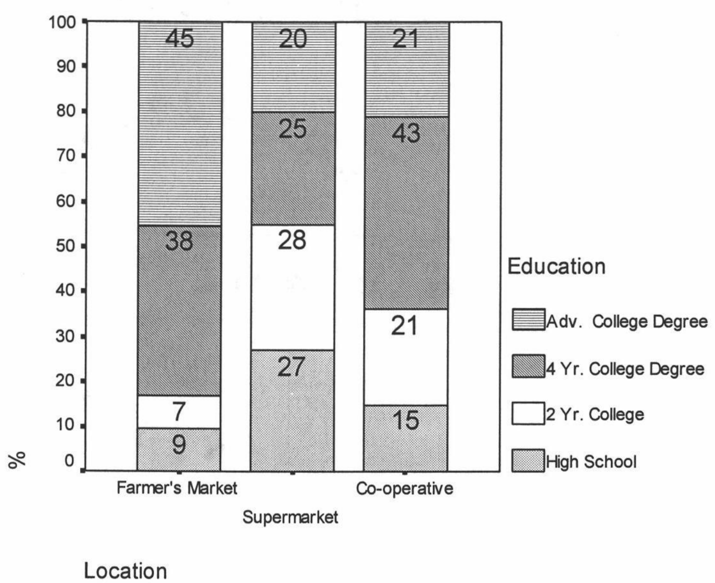


Figure 2: Education Level by venue



Appendix 3 (Continued): Cross tabulation of demographic variables by venue

Figure 3: Age by venue

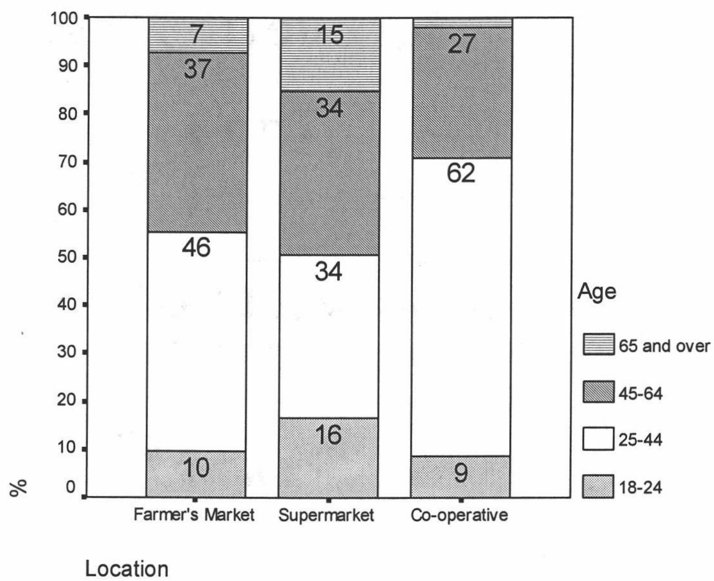
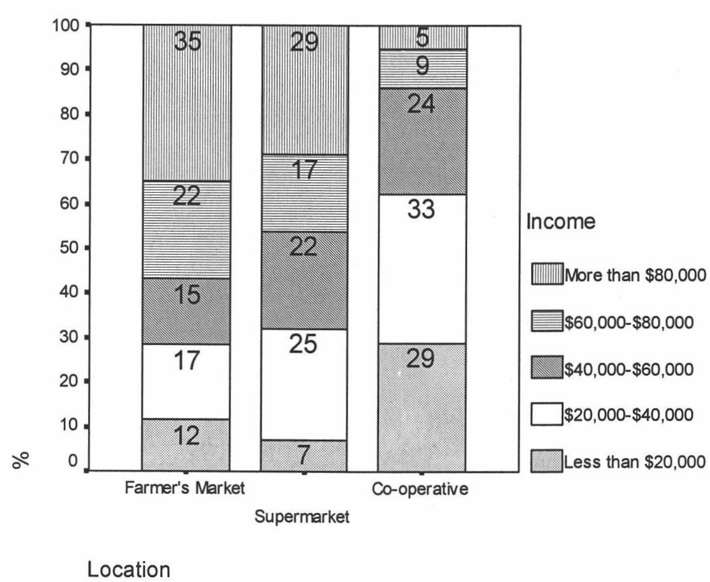


Figure 4: Income by venue



Appendix 4: Environmental questions and Frequencies

Table 2: Frequencies of environmental questions

	Never True		Rarely True		Sometimes True		Mostly True		Always True	
	Count	Row %	Count	Row %	Count	Row %	Count	Row %	Count	Row %
a) I recycle paper, cans or bottles	*2*	*.7%*	*5*	*1.8%*	*8*	*2.9%*	*80*	*28.6%*	*185*	*66.1%*
b) Do you look for the dolphin-safe label when you buy tuna?	*44*	*16.4 %*	*24*	*9.0%*	*24*	*9.0%*	*56*	*20.9%*	*120*	*44.8%*
c) I buy environmental friendly products	*25*	*9.0%*	*24*	*8.6%*	*72*	*25.8%*	*83*	*29.7%*	*75*	*26.9%*
d) I use rechargeable batteries	*60*	*21.4 %*	*61*	*21.8 %*	*75*	*26.8%*	*54*	*19.3%*	*30*	*10.7%*
e) I buy energy efficient light bulbs	*20*	*7.2%*	*38*	*13.6 %*	*62*	*22.2%*	*90*	*32.3%*	*69*	*24.7%*
f) I purchase recycled paper	*10*	*3.6%*	*29*	*10.5 %*	*86*	*31.0%*	*90*	*32.5%*	*62*	*22.4%*
g) I have switched product for environmental reasons	*27*	*9.7%*	*39*	*14.0 %*	*54*	*19.4%*	*87*	*31.3%*	*71*	*25.5%*
h) I have convinced others not to buy products harmful to the environment	*48*	*17.2 %*	*40*	*14.3 %*	*74*	*26.5%*	*64*	*22.9%*	*53*	*19.0%*
i) I will buy from companies that are ecologically irresponsible	*24*	*8.6%*	*36*	*12.9 %*	*79*	*28.2%*	*76*	*27.1%*	*65*	*23.2%*
j) I have tried very hard to reduce electricity use	*4*	*1.4%*	*13*	*4.7%*	*65*	*23.6%*	*101*	*36.6%*	*93*	*33.7%*
k) I have purchased products because they cause less pollution	*10*	*3.6%*	*21*	*7.6%*	*73*	*26.4%*	*119*	*43.1%*	*53*	*19.2%*
l) I try only to buy products that can be recycled	*7*	*2.5%*	*32*	*11.6 %*	*91*	*33.0%*	*110*	*39.9%*	*36*	*13.0%*
m) I do not buy household products that harm the environment	*7*	*2.5%*	*33*	*11.9 %*	*82*	*29.5%*	*108*	*38.8%*	*48*	*17.3%*

Appendix 5: Health Related questions

Table 3: Frequencies of health questions

	Never True		Rarely True		Sometimes True		Mostly True		Always True	
	Count	Row %	Count	Row %	Count	Row %	Count	Row %	Count	Row %
n) I avoid foods containing nitrites or preservatives	14	5.0%	37	13.3%	60	21.6%	111	39.9%	56	20.1 %
o) I am interested in information about my health	3	1.1%	9	3.2%	27	9.7%	75	27.1%	163	58.8 %
p) My daily diet is nutritionally balanced	4	1.5%	13	4.7%	55	20.0%	121	44.0%	82	29.8 %
q) I eat 5 servings of fruits and vegetables every day	10	3.6%	40	14.5%	72	26.1%	82	29.7%	72	26.1 %
r) I regularly participate in outdoor activities	5	1.8%	16	5.8%	62	22.4%	73	26.4%	121	43.7 %
s) I try to avoid stressful situations	9	3.3%	22	8.0%	79	28.6%	127	46.0%	39	14.1 %
t) I try to exercise at least 30 min a day, 3 days a week	12	4.3%	36	13.0%	51	18.5%	69	25.0%	108	39.1 %

	Strongly Disagree		Disagree		Neither Agree or Disagree		Agree		Strongly Agree	
	Count	Row %	Count	Row %	Count	Row %	Count	Row %	Count	Row %
a) I read more health related articles than I did 3 years ago	4	1.4%	24	8.6%	61	21.8 %	95	33.9 %	96	34.3%
b) I worry that there are harmful chemicals in my food	5	1.8%	24	8.6%	32	11.5 %	120	43.0 %	98	35.1%
c) I'm concerned about my drinking water quality	5	1.8%	11	3.9%	22	7.9%	96	34.4 %	145	52.0%
d) Good health takes active participation on my part	1	.4%	2	.7%	7	2.5%	74	26.7 %	193	69.7%
e) It is the doctor's job to keep me well	114	40.7 %	109	38.9 %	38	13.6 %	13	4.6%	6	2.1%
f) My health is outside my control	139	49.8 %	103	36.9 %	21	7.5%	11	3.9%	5	1.8%
g) I try to avoid high levels of cholesterol in my diet	8	2.9%	25	8.9%	62	22.1 %	123	43.9 %	62	22.1%
h) I exercise more than I did 3 years ago	15	5.4%	52	18.6 %	74	26.5 %	75	26.9 %	63	22.6%

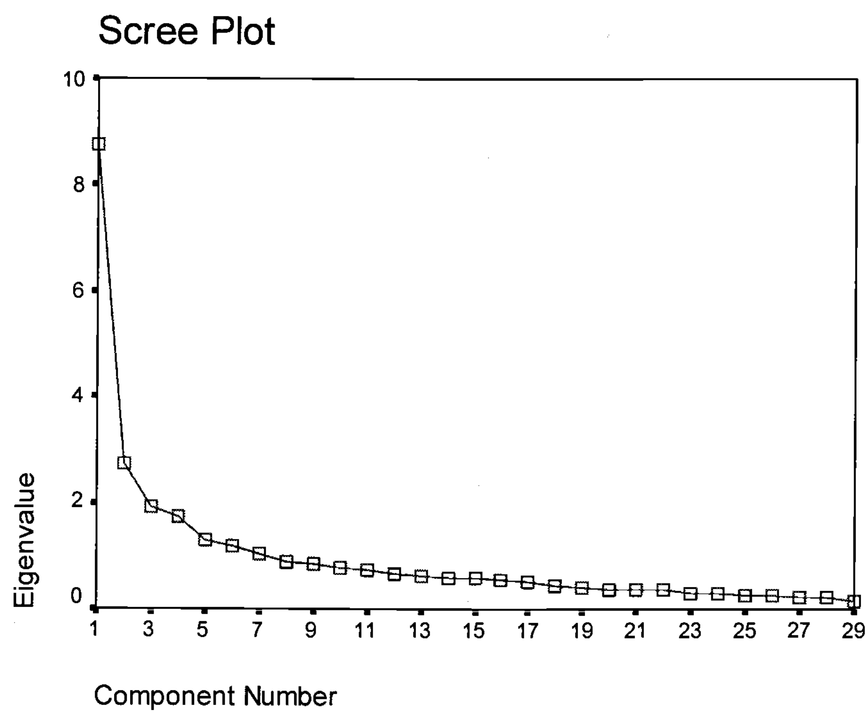
Appendix 6: Communalities and Scree Plot from the Factor Analysis

Table 4: Communalities

	Initial	Extraction
a) I recycle paper, cans or bottles	1.000	.451
b) Do you look for the dolphin-safe label when you buy tuna?	1.000	.485
c) I buy environmental friendly products	1.000	.739
d) I use rechargeable batteries	1.000	.723
e) I buy energy efficient light bulbs	1.000	.687
f) I purchase recycled paper	1.000	.640
g) I have switched product for environmental reasons	1.000	.761
h) I have convinced others not to buy products harmful to the environment	1.000	.668
i) I will buy from companies that are ecologically irresponsible	1.000	.638
j) I have tried very hard to reduce electricity use	1.000	.545
k) I have purchased products because they cause less pollution	1.000	.742
l) I try only to buy products that can be recycled	1.000	.742
m) I do not buy household products that harm the environment	1.000	.676
n) I avoid foods containing nitrites or preservatives	1.000	.573
o) I am interested in information about my health	1.000	.618
p) My daily diet is nutritionally balanced	1.000	.761
q) I eat 5 servings of fruits and vegetables every day	1.000	.651
r) I regularly participate in outdoor activities	1.000	.759
s) I try to avoid stressful situations	1.000	.342
t) I try to exercise at least 30 min a day, 3 days a week	1.000	.765
a) I read more health related articles than I did 3 years ago	1.000	.580
b) I worry that there are harmful chemicals in my food	1.000	.683
c) I'm concerned about my drinking water quality	1.000	.722
d) Good health takes active participation on my part	1.000	.442
e) It is the doctor's job to keep me well	1.000	.762
f) My health is outside my control	1.000	.727
g) I try to avoid high levels of cholesterol in my diet	1.000	.451
h) I exercise more than I did 3 years ago	1.000	.644
Free of GMO is important when buying fruits or vegetables	1.000	.604

Appendix 6 (Continued): Communalities and Scree Plot from the Factor Analysis

Figure 5: Scree Plot from the factor analysis



Appendix 7: Regression results for preference models on vegetables including free of GMO

Table 5: Regression results for preference model on vegetables including free of GMO

	Coeff.	Std.Err.	z-value	P-value
Constant**	4.300	1.683	2.555	0.011
Primary	0.257	0.535	0.481	0.631
Female	-0.480	0.409	-1.175	0.240
Age**	-0.061	0.015	-4.129	0.000
Kids	-0.542	0.443	-1.223	0.221
Education	0.220	0.220	0.998	0.318
Income	-0.079	0.061	-1.296	0.195
EPB**	0.550	0.252	2.184	0.029
ECRB*	-0.325	0.190	-1.710	0.087
Nutrition*	0.414	0.221	1.873	0.061
HESens	0.348	0.220	1.582	0.114
Fitness	0.190	0.193	0.981	0.326
PHResp	-0.181	0.194	-0.935	0.350
Member**	1.147	0.511	2.245	0.025
Price**	-0.989	0.361	-2.742	0.006
Appearance*	-0.583	0.331	-1.763	0.078
Free of GMO**	1.076	0.290	3.705	0.000

Dependent Var: VEGGIES *n*= 246

Measures of Fitness:
R^2_{ML} = 0.352 $R^2_{C\&U}$= 0.549

Frequencies of actual & predicted outcomes

Actual	Predicted 0	1	Total
0	61	29	90
1	18	138	156
Total	79	167	246
% of correct predictions[a]:			80.9%

Likelihood Ratio Test for all Bs=0[b]	
Log likelihood function	-96.97276
Restricted log likelihood	-161.5512
Chi-squared	129.1568
Degrees of freedom	16
Significance level	0.0000
Wald Test for all Bs=0[b]	
Wald Statistic	55.24019
Prob. from Chi-squared[16]	0.0000

*** Significant at the 0.05 level*
** Significant at the 0.10 level*

Table 6: Marginal effects for preference model on vegetables including free of GMO

Variable	Marginal at the means	Marginal on average[a]	Discrete 0-->1	Centered Δ1	Base Values[b]
Primary	0.061	0.027	0.063	---	1.000
Female	-0.115	-0.051	-0.107	---	1.000
Age**	-0.015	-0.006	---	-0.015	41.125
Kids	-0.129	-0.057	-0.134	---	0.000
Education	0.053	0.023	---	0.052	2.776
Income	-0.019	-0.008	---	-0.019	5.756
EPB**	0.131	0.058	---	0.131	0.000
ECRB*	-0.078	-0.034	---	-0.077	0.000
Nutrition*	0.099	0.044	---	0.099	0.000
HESens	0.083	0.037	---	0.083	0.000
Fitness	0.045	0.020	---	0.045	0.000
PHResp	-0.043	-0.019	---	-0.043	0.000
Member**	0.274	0.122	0.223	---	0.000
Price**	-0.236	-0.105	---	-0.232	2.382
Appearance*	-0.139	-0.062	---	-0.139	2.498
Free of GMO**	0.257	0.114	---	0.252	2.327

a: Base values- calculated at the values for each individual
b: Base values- continuous variables held at means, discrete held at mode