Two Stage Decision Model of Soy Food Consumption Behavior

Wanki Moon
Dept. of Agribusiness Economics
Southern Illinois University
Carbondale, IL 62901
E-mail: wmoon@siu.edu

Siva K. Balasubramanian
Dept. of Marketing
Southern Illinois University
Carbondale, IL 62901

Arbindra Rimal
Dept. of Agribusiness
Southwest Missouri State University
Springfield, Missouri 65804


Copyright 2004 by [authors(s)]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Introduction

The role of health information/knowledge in consumers’ decisions of food choices has attracted a great deal of attention from policy-makers and academics in recent years. On the one hand, knowledge of whether health information alters food consumption behavior is crucial for policy-makers (related to public health and nutrition) in designing effective strategies to reduce the incidences of chronic diseases. Academic researchers on the other hand have shown considerable scientific interests in quantifying the effect of health information/knowledge on health behaviors (e.g., Grossman, 1972; Grossman, 1976; Kenkel, 1991; Mooreman) in general and on food and dietary choices in particular (e.g., Variyam et al, 1999; Chern and Rickersten, 2002).

Our research examines whether consumers’ knowledge of health benefits of soy proteins affect their soy-based food consumption behavior. Clinical research in the past few years demonstrated that soy-based foods provide various health benefits relative to chronic diseases such as osteoporosis, heart disease, and cancer. For example, Messina and Barnes (1995) found that consuming soy-based foods had anticarcinogenic effects on breast and prostate cancers. More importantly, 25 grams of soy protein daily has been found to have a significant effect of lowering cholesterol level (Anderson, Johnstone, and Cook-Newell, 1995). In 1999, the Food and Drug Administration (FDA) permitted food companies to use health claims on soy-based foods that contain a minimum of 6.25 grams (i.e., one-fourth of the daily recommended level of 25 grams) of soy protein per serving. The approval was in response to a petition filed by Protein Technologies International (PTI), a Dupont Business, in May 1998.
Our research goal is to assess the impact of such medical research findings and regulatory confirmation on the development of the soy food market. Specifically, we examine whether consumers’ knowledge of the health benefits of soy foods impacts two decisions regarding (i) whether to participate in soy food market and (ii) how often (much) to consume (intensity of soy food consumption). We use a two-stage decision model (Heckman’s sample selection model) to accomplish the goal. An important advantage of the two-stage decision models is that, while the same set of explanatory variables may be included in the two equations, they are allowed to have differential effects between the two decisions. For example, several hypotheses can be proposed with respect to the possible effects of knowledge of health benefits of soy protein on the two-stage decision models: (i) knowledge of health benefits exert a strong and positive influence on both decisions, raising market penetration and promoting the consumption level of existing consumers; (ii) perceived health benefits significantly influence only the decision of whether to consume soy-foods, while having little impact on the decision of how much to consume; and (iii) perceived health benefits do not increase the probability of becoming a consumer of soy foods but motivate existing soy consumers to increase their level of consumption.

We conceptualize the link between perceived health benefits and soy consumption behavior using Lancaster’s characteristics and Fishbein’s multiattribute models. Other perceived attributes including taste, convenience, and cost of soy food are considered along with perceived health benefits. These attributes’ relative effects on soy consumption behavior are illustrated using simulation analysis. In addition to perceived attributes of soy food, the roles of general health-related factors such as health knowledge,
motivation, and awareness in determining soy consumption behavior are examined using mediation hypothesis.

**Previous Studies**

Evidence indicates that increased knowledge and awareness of the diet-chronic disease linkages motivates consumers to reduce their consumption of fat and cholesterol. For example, Brown and Schrader (1989) showed that consumer knowledge of the role of cholesterol in heart diseases caused a significant drop in shell egg consumption. Chern et al (1995) illustrated that accumulating health information and subsequent increases in awareness has resulted in significant increases in the consumption for vegetable oils and decreased consumption for butter and lard.

Other research assessed the impacts of marketing strategies using health claims and regulations against such strategies on the consumption of certain food products. FDA policy traditionally banned all health claims on food products. However, breakfast cereal industry began to use health claims in 1985 on their products in connection with the linkage between dietary fiber and lowered risks of colon cancer. This incidence prompted the FDA to review its policy concerning health claims and to suspend the ban. Ippolito and Mathios (1991) demonstrated that the marketing practices of using health claims led to significant increases in consumer knowledge of the fiber-cancer relationship, and resulted in the increased consumption of fiber cereal and in product innovation. Ippolito and Mathios (1995) also showed that the consumption of fats and saturated fats decreased faster after 1985 as compared to the period of 1977-1985.

Taken together, these studies illustrate that the private provision of diet-disease information was effective in altering consumer behavior over time and present evidence
that regulation impeded the flow of information in these particular markets. Since their
goal was to analyze aggregate behavioral change over time in connection with mounting
information of diet-disease linkages, it was not an issue of relevance to explain cross-
sectional (individual) differences in the access to health information and their impact on
behavior. It is reasonable, however, to anticipate that access to health information varies
across individuals potentially because of the differences in financial ability to search for
information and psychological factors (e.g., health motivation, health knowledge).
Several studies examined whether such cross-sectional differences in health knowledge
(information) led to behavioral discrepancies as measured by food choices or nutritional
intakes. For example, Variyam, Blaylock and Smallwood (1998) showed that greater
nutrition knowledge translated to significantly lower intake of dietary cholesterol.
Varyiam (1999) demonstrated that mothers’ health knowledge made a significant effect
on the dietary quality of their children.

**Conceptual Framework**

While traditional utility maximization framework is of little assistance in
understanding the role of various attributes in consumer demand for goods, Lancaster’s
characteristics model (1969) proposes that goods are not the direct objects of utility but it
is the characteristics of the goods from which utility is derived. Ladd and Suvannunt
(1976) extended Lancaster’s model in their consumer goods characteristics model
(CGCM) and derived two properties: (i) the price of a product is a sum of marginal
implicit value of attributes, and (ii) consumer demand for a product is influenced by the
level of attributes as well as prices and income. The first property gave rise to hedonic
price models as applied to decomposing the prices of beef and wheat into various
characteristics (Unnevehr and Bard, 1993; Espinosa and Goodwin, 1991). The second property was modified by Van Ravenswaay and Hoehn (1991) to focus on a single product and used by Baker and Crosbie (1993) to evaluate consumer preferences for food safety.

Our research uses the second property of the CGCM model as modified by van Ravenswaay and Hoehn (1991) as the theoretical model guiding our analysis of soy food consumption behavior in the U.S. Suppose that soy-based food product (\( Y \)) has a vector of salient characteristics \( C = \{ \text{taste, price, convenience, soy protein} \} \). Further, a vector of other products, \( A = (A_1, ..., A_l) \) are associated with a matrix of characteristics, \( a = [a_{ij}] \).

The key feature of this model is the recognition that consumers purchase the soy-based food and other products in order to derive non-market services such as satisfaction of appetite, balancing nutritional requirements and thereby maintaining health, or enjoyment of aesthetic pleasure. Equation (1) below shows that the quantity of such non-market services is determined by both market goods (\( Y \) and \( A \)) and level of attributes (\( C \) and \( a \)).

\[
S_i = S_i (Y, C, A, a),
\]

In consequence, appropriate arguments of consumers’ utility function are the non-market rather than market goods. Consumers’ utility maximization problem is represented by

\[
U = U (S_1, ..., S_l) \text{ subject to } m = P_1 * Y + P * A
\]

where \( P_1 \) and \( P \) represent the prices of the soy-based food and other products, respectively. Solving this constrained utility maximization problem, we obtain a demand function of soy food (\( Y \)),

\[
Y = Y (P_1, C, P, a, m)
\]
This equation indicates that consumers’ purchase decision for soy food is determined by economic variables (prices and income) and attributes of the soy food.

The model taking into account the level of characteristics reflects greater realism when compared to conventional models focusing on the demand for products themselves. Yet the current characteristic demand model potentially raises two issues with regard to the vector of attributes: (i) whether attributes are objectively measurable and (ii) whether consumers possess knowledge of such attributes.

The first issue is likely to be problematic particularly if the good in question is credence goods. Credence goods are those whose quality of attributes is rarely learned even after consumption, rendering it impossible to assess attributes objectively. With respect to the second issue, consumers may not possess knowledge of the attributes of goods, particularly with non-search properties. For example, soy-based foods were scientifically shown to have a nutritional attribute (isoflavones) that reduces the risks of coronary heart diseases. While the level of the isoflavones can be measured objectively, consumers may or may not have knowledge of the linkage between isoflavones and reduced risks of heart diseases. When consumers are not knowledgeable about the linkage, the level of isoflavones is of little relevance in determining consumer demand for the soy-based food product. Therefore, our research proposes that consumers’ knowledge/awareness of the attributes would be more appropriate in modeling soy food consumption behavior than the objective level of the attributes.

Fishbein’s multiattribute model offers a conceptual framework linking consumers’ perceptions of soy attributes to consumers’ behavior and attitude. Equation (5) shows the multiattribute model in a concise form.
(4) **Attitude** = Σ β Z

The model exhibits that attitude toward an object is determined by two factors: (i) the importance of the attributes associated with the object (β) and (ii) consumers’ beliefs (Z) that the object has the attributes. Under the premise that attitude toward soy foods is in correspondence to its consumption behavior, we replace C with Z in soy demand equation (3).

(5) \( Y = Y(P_1, Z, P, a_{ij}, m) \)

Now equation (5) represents the conceptual model resulting from the integration of characteristics demand model with consumers’ perceptions of such characteristics. The integrated model can be viewed as a special case of equation (3) that would be more appropriate when the study involves experience goods or goods with characteristics that are not widely recognized.

**Survey Design and Data**

We designed a survey instrument to elicit soy food consumption practices and various conceptual variables pertinent to analyzing consumer preferences for soy-based foods. The survey was administered in Sept. 2002 over the Web to members of the Ipsos-NPD panel. Of the 400,000 households participating in the Web panel, the Ipsos-NPD randomly selected 3,000 households appropriately stratified by geographic regions, income, education, and age to match with the U.S. census. The Ipsos-NPD sent e-mail letters to the 3,000 households prompting participants to the survey Web site. More than 1,400 households returned completed questionnaires, yielding a response rate of about 47%.
The instrument encompasses an array of questions intended to measure primarily four sets of variables: (i) perceived attributes associated with soy-based food products including price, taste, convenience of cooking and health benefits of soy protein; (ii) health-related variables including general nutritional knowledge, awareness and health motivation; (iii) socio-demographic profiles of respondents; and (iv) soy-based food consumption behavior as measured with monthly consumption frequency of Tofu, Soy veggie burgers, Soy milk, Soy Protein bars, Soy supplements, Soy cheese, and Meat Substitutes. Table 1 presents descriptions of the question items and summary statistics for the first three sets of variables that are used for our study.

**Perceived Attributes of Soy-based Foods** Consumer perceptions of health benefits of soy foods are measured with the following four items: soy foods (i) lower cholesterol level in your blood, (ii) act as an antioxidant, (iii) help retain bone mass, thereby reducing the risk of osteoporosis, and (iv) are good for women during menopause. Respondents were allowed to express their perceptions using a six-point agree/disagree scale. The four items were summed up and divided by four to construct an average index of perceived health benefits of soy proteins.

Three statements were posed to respondents to measure perceived convenience of cooking soy foods: (i) soy-based foods are convenient, (ii) recipes for soy-based foods are readily available, and (iii) I know how to prepare soy-based food products. An index of perceived convenience of soy-based foods was constructed by summing up these three items and dividing the sum by three, hence the index ranging from 1 (not convenient at all) to 6 (very much convenient). In addition, perceived price was measured with the
statement “soy foods are inexpensive” and perceived taste with “I like the taste of soy-based foods”.

General Health-related Variables Question items measuring general nutritional knowledge are drawn from Mooreman and Maulitch (1993). Respondents were asked to link eleven nutrients (i.e., Sodium, calcium, vitamin A, protein, vitamin C, iron, vitamin, carbohydrates, saturated fat, potassium and dietary fiber) to appropriate health consequences from the list of eleven items including high blood pressure, strong bones, healthy eyes, amino acids, anticancer power, oxygen, absorb calcium, conversion to sugar and fueling the body, cardiovascular disease, balancing sodium). An index of nutritional knowledge was constructed by adding up all correct answers for each respondent, hence the index ranging from a minimum score of 0 representing no nutritional knowledge at all to the maximum score of 11 representing maximum nutritional knowledge.

In addition, nutritional awareness was measured with the following eight items: (i) very concerned about the amount of salt in my diet, (ii) eat a lot of fresh fruit, (iii) eat a lot of fresh vegetables, (iv) actively trying to consume less fat in my diet, (v) actively trying to consume less cholesterol in my diet, (vi) very much concerned about linkages between diet and chronic diseases, (vii) very concerned about nutrition, (viii) concerned about getting enough calcium in my diet. Responses to the eight items were added up and divided by eight to obtain a constructed index of nutritional awareness. It ranges from a minimum score of 1 representing little concern about the effect of dietary choices on health to a maximum score of 6 representing a great deal of concern about health and nutrition.
Health motivation is measured with three items: (i) I read nutrition labels on food packages very carefully, (ii) I consulted a dietitian in the past year, and (iii) I changed my diet in the past years to reduce the risk of certain diseases. The inter-correlations of the three items were very high with correlation coefficients ranging from $\rho = 0.68$ to $\rho = 0.86$.

Socio-demographic Characteristics  Our survey instrument elicited demographic information including respondents’ age, education, income, household size, geographic regions, marital status and ethnic background. Nearly 91 percent of the respondents were whites, significantly under-representing Asians (1.8 percent) and African Americans (2.8) in our sample. About 56 % of respondents were from married households. Income category under $25,000 was 19.6 %, while categories of $25,000 \sim $49,999 and $50,000 and higher were 29 % and 51.4 %, respectively.

Soy Food Consumption Frequencies  Soy consumption behavior was measured with monthly frequency of serving products including Tofu, Soy veggie burgers, Soy milk, Soy Protein bars, Soy supplements, Soy cheese, and Meat Substitutes. Table 2 presents percentage of respondents reporting positive frequency of consumption for each of the soy food products and average monthly consumption frequency with and without zero reports included. Meat substitutes, Tofu, and Vege burger exhibited the greatest market penetration with more than 18 percent of respondents reporting positive level of consumption, followed by Soy milk, Soy supplements, and Soy cheese. Among soy-users, soy substitutes and soy milk was most frequently consumed with 14 and 10.35 times per month, respectively. When combined across the six categories, average consumption frequency was 5.5 times per month for the entire sample and 15.3 times for
the sub-sample of soy-users. More than 36 percent of total respondents reported that they consumed one of the five soy food products at least once per month.

**Two Stage Decision Models**

Our collected data indicate that market penetration rate of soy-based food products ranges from 8% ~ 19%, indicating that relatively small percentage of respondents is currently consuming soy-based food products on a regular basis. Therefore, it is critical to consider non-users when modeling soy consumption behaviors. Our research uses two-step decision approach that addresses such non-users as well as users. Two-step decision models explicitly incorporate participation decisions in an equation separate from consumption intensity decisions. Further, explanatory variables may have differential and even opposite effects in the two decision stages (Lin and Schmidt, 1983; Jones, 1989). Thanks to these merits, the two stage decision models have been widely applied to modeling various food consumption behaviors: aggregate food consumption (Haines et al, 1988), cheese (Gould, 1992), seafood products (Lin and Milon; 1993), cigarlette (Blisard and Blaylock, 1993), and meat (Burton et al, 1993; Ward, Moon and Sarina, 2002; Moon et al, 2002).

Two-step decision models refer commonly to (i) double-hurdle and (ii) first-hurdle dominance models. The double-hurdle model allows zero values to be either optimal continuous consumption decisions (i.e., corner solutions) or discrete desires not to participate in the market (i.e., non-users). In contrast, the first-hurdle dominance model differentiates between users and non-users: i.e., zero values indicate only non-users. Recalling that we measured soy consumption behavior as monthly consumption frequency aggregated over seven different soy-based food products, zero values from our
survey are likely to be indicative of non-users, pointing to the first-hurdle dominance as the appropriate model to fit our survey data. To derive likelihood function for the first-hurdle dominance model as a special case of double-hurdle model, we first describe the decision-making mechanism of the double-hurdle model below.

\[
Y_i = \begin{cases} 
Y_i^* & \text{if } Y_i^* > 0 \text{ and } D_i > 0 \\
0 & \text{otherwise}
\end{cases}
\]

where \( Y_i^* = X_i \beta + \varepsilon_i \)

\[
D_i = Z_i \theta + \nu_i
\]

Actual observed consumption of soy food is \( Y_i \) and \( Y_i^* \) is a latent variable representing optimal level of soy consumption. \( D_i \) is the zero-one discrete decision to participate in the soy market or not. A vector of explanatory variables (\( X_i \)) impacts the positive observations, while \( Z_i \) is a vector of variables explaining the decision to participate. Both \( D_i \) and \( Y_i^* \) must simultaneously be greater than zero to observe positive \( Y_i \) with the zeros indicating either optimal consumption decisions (corner solutions) or discrete decisions.

When the error terms in equation (1) follow a bivariate normal distribution with correlation coefficient \( \rho \), the double-hurdle decision-making framework is represented with the likelihood function shown below (Blundell and Meghir, 1987; Jones, 1989; Blaylock and Blisard, 1993);

\[
L = \prod_0 [1 - \Pr (D_i = 1) \Pr (Y_i^* > 0 | D_i = 1)] \\
\prod_1 \Pr (D_i = 1) \Pr (Y_i^* > 0 | D_i = 1) \mathcal{N} (Y_i^* | Y_i^* > 0, D_i = 1)]
\]

A special case of the double-hurdle arises when there are no corner solutions, with all zero observations signaling that the household does not consume soy food products (Jones, 1989; Blaylock and Blisard, 1993). In this case, the zero observations do not affect the parameters in the second-stage consumption equation, thus leading to specifications commonly noted as the first-hurdle dominance model. The statistical
counterpart to the first-hurdle dominance model is Heckman’s sample selection model which differs from equation (2) in that only $D_i$ needs to be greater than 0 to observe positive $Y_i$. In this model, once the consumer decides to participate in the soy market, the observed soy consumption is always positive. The first-hurdle dominance decision rule for soy consumption is represented with the likelihood;

$$L = \prod_0 [1 - \Pr (D_i = 1)] \prod_1 \Pr (D_i = 1) N [Y_i^* | D_i = 1]$$

Equations (7) and (8) differ in two components: (i) $\Pr (Y_i^* > 0 | D_i = 1)$ is equal to one and (ii) $N [Y_i^* | Y_i^* > 0, D_i = 1]$ is reduced to $N [Y_i^* | D_i = 1]$. These two differences clearly establish that once deciding to enter the soy market, no corner solutions are observed. The log likelihood function corresponding to equation (3) is shown below (MacKinnon and Davidson, 1993),

$$\ln L = \sum_0 \ln (1 - \Phi (Z_i \theta)) + \sum_1 \ln \{\phi (Y_i - X_i \beta)/\sigma\}
\sum_1 \ln \Phi \{Z_i \theta + \rho (Y_i - X_i \beta)/\sigma\}/(1 - \rho^2)^{0.5}\}
$$

When the correlation coefficient in equation (4) is zero, the model can be separated into the probit for the first-stage and OLS for the consumption equation, yielding a complete dominance model.

**Model Specification**

While $X$ and $Z$ in the sample selection model may include different sets of variables, there is little reason to differentiate between the two vectors in this particular case of modeling soy consumption behavior. Therefore, $X$ and $Z$ will include an identical set of explanatory variables. Yet, the variables in the two vectors are allowed to have differential impacts between the two stages. Table 1 shows brief description of explanatory variables included in the sample selection model and summary statistics.
Drawing on our conceptual model as represented by equation (5), the vector (X) includes consumer perceptions of price, taste, convenience, and health benefits of soy proteins.

Moreover, our research recognizes that whether consumers perceive health benefits from soy foods is an endogenous construct that could be shaped by other health-related factors. In consequence, we hypothesize that perceived health benefits are determined by (i) how much consumers are knowledgeable about the linkage between diet and diseases (Health Knowledge), (ii) how motivated consumers are in seeking health (Health Motivation), or (iii) Nutritional Awareness as well as by demographic profiles including age, education, ethnic background, income and gender. Yet, these variables can affect consumers’ soy food consumption decisions directly as well. To streamline these interdependent relationships among variables in our empirical models, we propose that perceived health benefits mediate the effects of health knowledge, awareness and motivation on soy market participation and consumption intensity decisions.

Figure 1 depicts such interdependent relationships among perceived attributes, general health-related variables, and soy behavior. Soy behavior is represented by two endogenous decisions of whether to participate in soy market and how much to consume. These two decisions are shown to be determined by perceived attributes and general health-related variables. Further, the figure shows that perceived health benefits can mediate the impact of general health knowledge, motivation, and awareness on soy behavior.

We test this mediating hypothesis using three criteria suggested by Baron and Kenny (1986): (i) perceived health benefits have a statistically significant impact on soy
food consumption decisions, (ii) health knowledge, motivation and nutritional awareness significantly influences soy food consumption decisions, (iii) these health-related variables exert significant influence on perceived health benefits, and (iv) the effects of these health-related variables are diminished or no longer significant when estimated along with perceived health benefits. These criteria require that four separate sets of regression models be estimated,

\[
\begin{align*}
\text{(a1) soy market participation decision} &= f(\text{perceptions about health benefits, price, taste and convenience}) \\
\text{(a2) soy consumption intensity} &= f(\text{perceptions about health benefits, price, taste and convenience},) \\
\text{(b1) soy market participation decision} &= f(\text{health knowledge and motivation and nutritional behavior, age, education, income ethnic background, marital status}) \\
\text{(b2) soy consumption intensity} &= f(\text{health knowledge and motivation and nutritional behavior, age, education, income ethnic background, marital status}) \\
\text{(c) Perceived health benefits} &= f(\text{health knowledge and motivation and nutritional behavior, age, education, income ethnic background, marital status}) \\
\text{(d1) soy market participation decision} &= f(\text{perceived health benefits, price, taste and convenience, health knowledge and motivation and nutritional behavior, age, education, income ethnic background, marital status}) \\
\text{(d2) soy consumption intensity} &= f(\text{perceived health benefits, price, taste and convenience, health knowledge and motivation and nutritional behavior, age, education, income ethnic background, marital status})
\end{align*}
\]

Equation (a) tests the first criterion, while equations (b) and (c) test the second and third criteria, respectively. Comparing estimated parameters across models (a) through (d) will permit us to assess whether perceived health benefits mediate the effects of other health-related variables on soy consumption decisions.

In summary, our model specification reflects that perceived health benefits specific to soy foods impact soy food consumption decisions, while other health-related
variables may shape such perceptions, thereby indirectly impacting soy food consumption decisions.

**Results**

Sample selection model was estimated for models (a), (b) and (c) using Heckman’s two-step estimation procedures (Greene, 1998). Perceived health benefit equation (c) was estimated by ordinary least square method. Table 3 shows estimation results for models (a) through (d). Model (a) links perceived attributes to soy consumption decisions. Estimated results show that perceptions about health benefits, taste, and convenience have significant impact on the probability of consuming soy-based food products (i.e., decision of whether to consume) as well as on the decision of how often to consume. For example, respondents who perceive positive health benefits from soy foods are more likely to participate in the soy market and consume soy foods more often when compared to those who do not perceive health benefits. Further, respondents who perceive soy foods as tasty and convenient were significantly more likely to participate in the soy market and consume soy foods more often when compared to those who do not have such perceptions. Perceived price, however, was not important in explaining both decisions.

Model (b) relates health knowledge, motivation and awareness and socio-demographic characteristics to soy consumption decisions. Estimated results display that these three nutrition-related variables significantly affect both the likelihood of participating in the soy market and consumption intensity (except for Health Knowledge which did not have a statistically significant impact in the consumption frequency equation). The more motivated to seek healthy diets and more aware of the importance
of diets in their health, the more likely to participate and consume soy foods. Socio-demographic characteristics also make significant differences in the two decisions. In particular, respondents with higher education were more likely to participate in the soy market and consume soy foods more often. Racial background was statistically significant: i.e., whites were less likely to participate in the soy market when compared to other races including African-American, Asians, and Hispanics.

Model (c) hypothesizes that whether consumers know of health benefits of soy foods would be determined by their health knowledge, motivation and awareness along with socio-demographic characteristics. Estimated results clearly show that perceived health benefits are shaped by these three health-related variables. Respondents with greater health knowledge, motivation and awareness are more likely to perceive health benefits of soy foods. Perceived health benefits also differ across socio-demographic profiles. Respondents with higher education perceived greater health benefits from soy foods. In addition, female and older respondents were more likely to see health benefits from soy foods.

Model (d) incorporates all sets of variables that are hypothesized to explain soy food consumption behavior. We compare the results of model (d) with models (a) and (b) in order to offer insight into whether perceived health benefits of soy foods mediate the effects of other general health-related variables on soy food consumption behavior. Estimated results show that health knowledge has little impact on the probability of participating in the soy market and consumption frequency. This result, when combined with results from (a), (b) and (c), indicates that its effect on soy food consumption behavior is mediated by perceived health benefits. In contrast, health motivation and
awareness are still significant, although the size of impact is reduced. This result suggests that health motivation and awareness exert a direct influence on soy consumption behavior, as well as indirectly via perceived health benefits. That is, respondents with greater health motivation and awareness are more likely to perceive health benefits from soy foods and, at the same time, more likely to participate in the soy market and consume soy foods. Education also impacts soy consumption decisions indirectly via positive impact on perceived health benefits as well as directly affecting soy market participation and consumption intensity.

**Simulating the Impact of Perceived Attributes**

The analysis of sample selection model estimates empirically establishes that perceived health benefits, taste, and convenience play a significant role in shaping consumers’ behavior with respect to soy foods. In particular, the estimated impact of perceived health benefits confirms that the growing role of health concerns in food choices in general is also pertinent in the soy market. A question remains as to the role of the perceived health benefits relative to other attributes such as taste or convenience. We conduct simulation analysis to shed light on this question using parameter estimates from Model (a).

Figure 2 shows the impacts of the three attributes on the probability of participating in soy market. The probability increases from 20 percent to 50 percent as the index of perceived health benefits changes from “strongly disagree” (1) to “strongly agree” (7). Yet, the figure clearly shows that perceived taste and convenience had stronger impacts on the probability of participating in soy market when compared to perceived health benefit. In particular, the probability decreases drastically from 80
percent to less than 10 percent when the index of perceived taste changes from 7 (agree that soy food is tasty) to 1 (disagree). When consumers’ perceptions about convenience of soy food changes from 7 (agree that soy food is convenient) to 1 (disagree), the probability diminishes from about 58 percent to 12 percent.

Using second-stage parameter estimates, figure 3 illustrates the relative impacts of the three attributes on the behaviors of current soy-users as measured by frequency of consumption. Perceived taste and convenience are shown to exert comparable impact on the frequency of soy consumption, while being considerably larger when compared to perceived health benefits. When respondents strongly agree that soy food is tasty or convenient, their simulated consumption frequency reaches nearly 40 times a month. Current soy-users consume on average about 23 times a month when being assured that soy food offers health benefits relative to certain chronic diseases.

Conclusions/Implications

This study examined the role of soy health benefits in consumers’ soy consumption decision. Given the large number of respondents who reported no consumption of soy products per month, it was important to model the decision of whether or not to participate in soy market separately from the consumption intensity decision. Estimation results demonstrate that knowledge of health benefits affects both the likelihood of participation and consumption intensity. That is, consumers with higher soy health knowledge are more likely to enter soy food market and consume soy products more often when compared to those with less soy knowledge. This result is generally in line with time-series findings (Brown and Schrader, 1990; Chern et al, 1995; Ippolito and Mathios, 1991; Ippolito and Mathios, 1995) and cross-sectional evidence (Variyam,
Blaylock and Smallwood, 1998; Variyam et al, 1998) regarding the role of health knowledge in food/dietary choices.

Simulation analysis revealed that perceived taste and convenience of soy foods had greater impact than perceived health knowledge on both participation and consumption intensity: i.e., respondents who do not consider soy foods as appetizing and convenient to use were significantly less likely to be soy-users and, even if they became soy-users, they were likely to use soy foods less frequently. In other words, although health benefits of soy foods motivates consumers to participate and consume more soy foods, perceived unappetizing taste and lack of knowledge (convenience) of cooking prevent consumers from eating soy foods to a greater proportion than the favorable motivation due to health benefits. These results appear to expressively depict the current soy market which is growing obviously thanks to health benefits of soy proteins, but still majority of U.S. consumers are staying out of the market because of soy’s unattractive taste and inconvenience.
References


Table 1. Soy-based food consumption behavior: participation and frequency

<table>
<thead>
<tr>
<th>Soy Behavior</th>
<th>Percent of positive consumption (%)</th>
<th>Mean consumption frequency (all obs.)</th>
<th>Mean consumption frequency (positive obs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tofu</td>
<td>18.63</td>
<td>0.79</td>
<td>4.17</td>
</tr>
<tr>
<td>Soy Milk</td>
<td>12.53</td>
<td>1.29</td>
<td>10.35</td>
</tr>
<tr>
<td>Vege Burger</td>
<td>18.48</td>
<td>0.70</td>
<td>3.76</td>
</tr>
<tr>
<td>Meat Substitutes</td>
<td>18.9</td>
<td>1.14</td>
<td>6.05</td>
</tr>
<tr>
<td>Soy supplements</td>
<td>7.98</td>
<td>1.12</td>
<td>14.09</td>
</tr>
<tr>
<td>Soy cheese</td>
<td>6.32</td>
<td>0.53</td>
<td>8.35</td>
</tr>
<tr>
<td>Total</td>
<td>36.3</td>
<td>5.57</td>
<td>15.31</td>
</tr>
</tbody>
</table>

Figure 1. Mediating Role of Soy Health Benefits in Determining Soy Behavior
Table 2. Description and summary statistics of variables in model estimation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceived Soy Attributes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Health Benefits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowering cholesterol</td>
<td>Soy foods lower cholesterol in your blood</td>
<td>4.75</td>
<td>1.32</td>
</tr>
<tr>
<td>Antioxidant</td>
<td>Soy foods act as an antioxidant</td>
<td>4.55</td>
<td>1.22</td>
</tr>
<tr>
<td>Bone mass</td>
<td>Soy foods help retain bone mass</td>
<td>4.35</td>
<td>1.32</td>
</tr>
<tr>
<td>Menopause</td>
<td>Soy foods are good for women during menopause</td>
<td>4.74</td>
<td>1.36</td>
</tr>
<tr>
<td>(ii) Convenience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenient</td>
<td>Soy foods are convenient</td>
<td>3.59</td>
<td>1.45</td>
</tr>
<tr>
<td>Recipes</td>
<td>Recipes of soy foods are readily available</td>
<td>3.74</td>
<td>1.59</td>
</tr>
<tr>
<td>Preparation</td>
<td>I know how to prepare soy foods</td>
<td>2.64</td>
<td>1.77</td>
</tr>
<tr>
<td>(iii) Taste</td>
<td>I like the taste of soy foods</td>
<td>3.14</td>
<td>1.73</td>
</tr>
<tr>
<td>(iv) Price</td>
<td>Soy foods are inexpensive</td>
<td>3.24</td>
<td>1.48</td>
</tr>
<tr>
<td><strong>Health Knowledge</strong></td>
<td>Aggregation of correct answers to 11 nutritional questions</td>
<td>8.29</td>
<td>2.65</td>
</tr>
<tr>
<td><strong>Health Motivation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrition label</td>
<td>I read nutritional labels very carefully</td>
<td>4.26</td>
<td>1.84</td>
</tr>
<tr>
<td>Dietitian</td>
<td>I consulted a dietitian in the past year</td>
<td>1.71</td>
<td>1.52</td>
</tr>
<tr>
<td>Changing Diet</td>
<td>I changed my diet to reduce the risk of certain diseases</td>
<td>3.29</td>
<td>1.94</td>
</tr>
<tr>
<td><strong>Health Awareness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fresh fruit</td>
<td>I eat a lot of fresh fruits</td>
<td>4.16</td>
<td>1.67</td>
</tr>
<tr>
<td>Fresh vegetables</td>
<td>I eat a lot of fresh fruits</td>
<td>4.39</td>
<td>1.61</td>
</tr>
<tr>
<td>Less fat</td>
<td>I actively try to consume less fat in my diet</td>
<td>4.59</td>
<td>1.67</td>
</tr>
<tr>
<td>Less cholesterol</td>
<td>I actively try to consume less cholesterol</td>
<td>4.32</td>
<td>1.72</td>
</tr>
<tr>
<td>Chronic diseases</td>
<td>I am very concerned about the role of diet in chronic diseases</td>
<td>4.57</td>
<td>1.63</td>
</tr>
<tr>
<td><strong>Socio-Demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>1=less than $5,000; 25=$250,000 or more</td>
<td>11.4</td>
<td>5.6</td>
</tr>
<tr>
<td>Age</td>
<td>In years</td>
<td>45.3</td>
<td>12.54</td>
</tr>
<tr>
<td>Education</td>
<td>1=Grade, 2=Some high, 3=High graduated, 4=Some college, 5=2 year college, 6=4 year college, 7=Some post graduate, 8=Post graduate degree</td>
<td>5.0</td>
<td>1.63</td>
</tr>
<tr>
<td>Gender</td>
<td>1 if Male; 0 otherwise</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Ethnic Background</td>
<td>1 if Whites; 0 otherwise</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3: Soyfood Consumption Behavior: Estimated parameters of the empirical models

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model I</th>
<th>Model II</th>
<th>Perceived Health Benefits (Model III)</th>
<th>Model IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated Parameter</td>
<td>T-ratio</td>
<td>Estimated Parameter</td>
<td>T-ratio</td>
</tr>
<tr>
<td><strong>Participation:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Benefit</td>
<td>0.147*</td>
<td>3.816</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Taste</td>
<td>0.362*</td>
<td>12.572</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cost</td>
<td>-0.017</td>
<td>-0.590</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Convenience Health</td>
<td>0.215*</td>
<td>5.415</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge</td>
<td>-</td>
<td>-</td>
<td>0.062*</td>
<td>5.763</td>
</tr>
<tr>
<td>Motivation</td>
<td>-</td>
<td>-</td>
<td>0.051*</td>
<td>6.398</td>
</tr>
<tr>
<td>Awareness</td>
<td>-</td>
<td>-</td>
<td>0.025*</td>
<td>6.919</td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>-</td>
<td>-0.003</td>
<td>-1.227</td>
</tr>
<tr>
<td>Education</td>
<td>-</td>
<td>-</td>
<td>0.106*</td>
<td>6.303</td>
</tr>
<tr>
<td>Income</td>
<td>-</td>
<td>-</td>
<td>0.045</td>
<td>1.259</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>-</td>
<td>-</td>
<td>-0.545*</td>
<td>-5.533</td>
</tr>
<tr>
<td>Log likelihood function</td>
<td>-710.015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted log likelihood</td>
<td>-960.374</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-squared</td>
<td>500.718*</td>
<td></td>
<td>373.323*</td>
<td></td>
</tr>
<tr>
<td><strong>Purchase:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONE</td>
<td>-104.179*</td>
<td>-4.047</td>
<td>-92.870</td>
<td>-1.435</td>
</tr>
<tr>
<td>Health Benefit</td>
<td>4.844*</td>
<td>3.465</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Taste</td>
<td>8.996*</td>
<td>3.852</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cost</td>
<td>-1.136</td>
<td>-1.333</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Convenience Health</td>
<td>8.243*</td>
<td>5.100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Knowledge</td>
<td>-</td>
<td>-</td>
<td>1.463</td>
<td>1.178</td>
</tr>
<tr>
<td>Motivation</td>
<td>-</td>
<td>-</td>
<td>1.747**</td>
<td>1.828</td>
</tr>
<tr>
<td>Awareness</td>
<td>-</td>
<td>-</td>
<td>1.070*</td>
<td>2.228</td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>-</td>
<td>-0.090</td>
<td>-0.927</td>
</tr>
<tr>
<td>Education</td>
<td>-</td>
<td>-</td>
<td>3.113</td>
<td>1.553</td>
</tr>
<tr>
<td>Income</td>
<td>-</td>
<td>-</td>
<td>-0.969</td>
<td>-0.602</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>-</td>
<td>-</td>
<td>-12.667</td>
<td>-1.292</td>
</tr>
<tr>
<td>LAMBDA</td>
<td>35.199*</td>
<td>3.510</td>
<td>39.706</td>
<td>1.473</td>
</tr>
<tr>
<td>F-statistics</td>
<td>31.060*</td>
<td></td>
<td>6.300*</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.224</td>
<td>0.051</td>
<td>0.188</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at ∀<0.05; **Significant at ∀<0.10
Figure 2. Simulated Impact of Perceived Attributes on Probability of Participation.

Figure 3. Simulated Impact of Perceived Attributes on Frequency of Consumption.