Discounting Spotted Apples: Investigating Consumers' Willingness to Accept Cosmetic Damage in an Organic Product

Chengyan Yue, Frode Alfnes, and Helen H. Jensen

Working Paper 06-WP 436
November 2006

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Chengyan Yue is a graduate assistant in the Department of Economics and Center for Agricultural and Rural Development (CARD), Iowa State University. Frode Alfnes is a postdoctoral fellow, Department of Economics and Resource Management, Norwegian University of Life Sciences. Helen Jensen is a professor in the Department of Economics and CARD, Iowa State University.

The authors thank Daren Mueller, Ariun Ishdorj, Chun-Fu Chen, and S. Patricia Batres-Marquez for help with the experiment. Frode Alfnes thanks the Research Council of Norway, grant no. 159523/110, for financial support. The USDA Risk Management Agency provided partial funding for the research.

Questions or comments about the contents of this paper should be directed to Helen H. Jensen, 578 Heady Hall, Iowa State University, Ames, IA 50011-1070 (515) 294-6253; Fax: (515) 294-6336; E-mail: hhjensen@iastate.edu.

Abstract

Organic producers have limited methods of avoiding plant diseases that result in cosmetic damage to produce. Therefore, the appearance of organic produce is often less than perfect. We use an experimental auction to investigate how cosmetic damage affects consumers' willingness to pay for organic apples. We find that 75% of the participants are willing to pay more for organic than for conventional apples given identical appearance. However, at the first sight of any imperfection in the appearance of the organic apples, this segment is significantly reduced. Furthermore, we find that there is a significant effect of interaction between cosmetic damage and product methods. Even though most consumers say they buy organic products to avoid pesticides, we find that cosmetic damage has a larger impact on the willingness to pay for organic apples than for conventional apples.

Keywords: appearance, apples, experimental auctions, organic, willingness to pay.

1. Introduction

Until recently, fresh food products such as apples were provided to markets as generic products. Now, however, these products are differentiated by brand, variety, origin, and appearance, as well as by the companies' production and processing methods. Consumers are often willing to pay large price premiums for products with the right attributes. As a result, product quality and differentiation have become increasingly important to the producer.

Empirical estimates of price variation due to quality factors date back at least to Waugh's seminal study of quality factors affecting vegetable prices (Waugh, 1928). One of the most important quality factors is appearance. Appearance includes the intrinsic attributes of color, texture, size, uniformity and other visible differences. Several recent studies consider how appearance affects consumers' preference for food products; see Acebron and Dopico (2000) for beef; Alfnes *et al.* (2006) for salmon; and Wei *et al.* (2003) for mandarin oranges. Credence attributes (Darby and Karni, 1973) are other quality factors valued by consumers. For fruits, organic production methods are a credence attribute since consumers cannot identify a product's method of production through their normal use of the product but have to trust the labeling.

Most previous studies investigating consumer preference for organic foods assume that the organic products are similar in appearance to their conventionally produced counterparts (Blend and van Ravenswaay, 1999; Loureiro, McCluskey, and Mittlehammer, 2001; and Larue *et al.,* 2004). Studies that focus on the effect of cosmetic problems find that consumers discount products with cosmetic damage. In a retail setting, Thompson and Kidwell (1998) found that the more cosmetic defects there were in organic produce, the less likely were shoppers to buy the organic produce. Experimental-based results from Roosen *et al.* (1998) show that if cosmetic attributes are the same, consumers tend to pay a positive premium for nonuse of pesticides. However, if the nonuse of pesticides results in products with reduced cosmetic quality, fewer consumers prefer nonuse of pesticides. Baker (1999) conducted a survey involving consumer preferences for food safety attributes in fresh apples (specifically, reduced or no pesticide use) and took account of the damage level on red delicious apples using pictures. Using clustering techniques, he found that cosmetic damage was most important to consumers with higher income. The three studies all find a positive effect from organic production (or nonuse of pesticides) and a negative effect of cosmetic damage. However, less well understood is the nature of the trade-off—whether the measured response to damage is sensitive to the production method, and what the effect is of underlying consumer attitudes about production method, environmental issues, and other quality attributes.

In this paper, we use a fourth-price sealed-bid auction to elicit consumer willingness to pay (WTP) for organic and conventional apples with different levels of blemish. In contrast to the consumer studies discussed earlier, we use an experimental design that allows not only the estimation of the main effects of production method and cosmetic damage but also the interaction effects between the two. What is new to the experimental auction mechanism used here is the individual drawing of a binding alternative. This allows us to combine the positive features of the incentive-compatible fourth-price auction with another feature imperative in a WTP study of products that are heterogeneous in so many ways, such as apples: the products the participants evaluated were the exact same products they would buy. The individual drawing of a binding alternative ensured that there was never more than one buyer of each alternative.

A principal component factor analysis and random effect models are used in the analysis of how the WTP for apples is affected by quality attributes (conventional versus organic production methods, degree of blemish, and their interaction), as well as interactions among consumers' stated attitudes toward specific quality attributes (food safety concern, environmental concern, tolerance of pesticides, etc.), production method, degree of blemish, and consumers' socio-demographic characteristics. Specifically, we investigate the premium for organic apples, the discount for various levels of cosmetic damage, how cosmetic damage affects consumers' WTP for both organic and conventional apples, and how attitude and socio-demographic variables affect these premiums.

2. Market Experiment

The experiment had a within-subject design with two production methods (organic and conventional), four appearance levels (degrees of blemish), and two elicitation methods (hypothetical and real auctions). In addition, we collected numerous socio-demographic and attitude measures.

2.1 Products

The products used for this experiment were 3-pound bags of golden delicious apples. Apples were obtained from commercial sources and from university farm orchards. Prior to the experiment, the apples were sorted according to their production method and appearance. The production methods included both conventional and organic methods. For the appearance, the apples were then sorted by the level of surface blotches (cosmetic damage). The blotches were caused by plant diseases and syndromes, namely, sooty blotch fungi and russeting, that led to changes that were strictly cosmetic and presented no harm to humans or to the taste of apples.

The conventional apples were sorted into four grades: SpotA apples were those without blotches; SpotB apples were those with about 3% blotch coverage; SpotC apples were those with about 5% blotch coverage; and SpotD were those with about 9% blotch coverage. The classification of apples was done with assistance from staff with training in plant pathology. Because of the lack of variation in their appearance, the organic apples were only sorted into two grades: SpotA, apples without blotches; and SpotB, apples with 3% blotch coverage. All of the sorted apples were packed into clear bags. We will, hereafter, refer to organic SpotA apples as Organic A, and conventional SpotA as Conventional A, and so on.

In the experiment, 12 bags of apples were placed on a large table for visual inspection. The apples were labeled as organic or conventional but were not labeled with the appearance grade. Instead, participants examined the appearance of the apples and made bids based on their own observations. Each alternative in the experiment had one specific bag of apples, and several of the alternatives had the same characteristics with respect to production method and cosmetic damage. Except for the aforementioned heterogeneous appearance, each bag contained apples that were as homogeneous as possible in other characteristics, such as number, size, and weight.

We also ran treatments in which we used pictures of apples. The apples in the pictures were 3-pound piles, sorted by appearance. In the picture treatments, we had four levels of cosmetic damage for both the organic and the conventional apples.

2.2 Experimental Procedure

We conducted fourth-price sealed-bid auctions with simultaneous bidding on 12 alternatives. A fourth-price sealed-bid auction is an auction in which the bidders submit sealed bids and the price is set equal to the fourth-highest bid; the winners are those who have bid more than the price. Vickrey (1961) showed that, in such an auction in which the price equals the first-rejected bid, it is a weakly dominant strategy for people to bid their true WTP for the offered goods. People have an incentive to truthfully reveal their private preferences because the auction separates what they say from what they pay. Consumers who underbid risk foregoing a profitable purchase, whereas consumers who overbid risk making an unprofitable purchase. In the last 15 years, experimental auctions have been used to elicit WTP for a wide variety of food quality attributes (see, e.g., Alfnes and Rickertsen, 2003; Lusk, Feldkamp, and Schroeder, 2004; Lusk *et al.*, 2004; Melton *et al.*, 1996; Roosen *et al.,* 1998; Rozan, Stenger, and Willinger, 2004; Umberger and Feuz, 2004).

Recently, several studies have used a uniform *n*th price auction such as ours to elicit WTP for food quality characteristics. See, for example, Umberger and Feuz (2004) for an application of a fourth-price sealed-bid auction, and Lusk *et al.* (2004) for an application of a fifth-price sealed-bid auction. Compared with the frequently used second-price auction, the fourth-price and other uniform *n*th price auctions have several benefits. First, if there are multiple winners, a winning position does not lead to an exclusive winner and any auction-winning utilities not associated with the product are reduced. Second, in a fourth-price auction with seven or more participants there is a smaller difference between the median participant's valuation of the product and the price. Therefore, a bid that differs from a participant's WTP is more likely to have real economic consequences. Third, with repeated trials, extreme outliers are less likely to affect the price information that the participants receive during the multi-trial experiments.

After the auction, each participant randomly drew his or her exclusive binding alternative. The drawing was done without replacement; only one participant could draw each of the alternatives as his or her binding alternative. For this to be possible, the number of alternatives had to be higher than or equal to the number of participants in each session. The price of an alternative was equal to the fourth-highest bid for that alternative. If the participants had bid more than the price for their binding alternative they had to buy the alternative. This winning restriction allowed us to combine the attractive features of the uniform-price auction (discussed earlier) with another feature that we felt was imperative in a WTP study of appearance of a heterogeneous product such as apples: the products they evaluated were the exact same products they would buy.

At the beginning of each session, the participants were given a folder containing US$20, a consent document, and a questionnaire. There were a total of eight sessions. In six of the eight sessions, we first conducted a hypothetical auction in which the apples were represented by pictures. We asked participants to examine carefully the apples in the pictures before they made their hypothetical bids. After the hypothetical auctions, we replaced the pictures with actual apples and ran one trial with a non-hypothetical auction. In the last two sessions, we did not run a hypothetical auction. Instead, we ran two trials with real auctions.[1] To avoid income and substitution effects, we randomly drew which of the two real auction trials was to be binding and then drew individual binding products.[2]

The participants walked around the table and placed their bids on their bidding forms as they studied each alternative. The participants were not allowed to communicate with each other during the bidding process. To reduce any systematic ordering effects, the participants could start at any of the 12 alternatives on the table. In the picture treatments, we had three pictures from each of the four categories of cosmetic appearance. Half of the pictures were labeled as organic. In the second half of the sessions, the other half of the pictures were labeled as organic. Thus, all the pictures were labeled as organic in half of the sessions and as conventional in the other half. This was done to reduce any unforeseen effects from small differences in the pictures. When using real products (actual apples) we had only SpotA and SpotB organic apples but we had all four categories of conventional apples.

2.3 Experimental Subjects

The experiment was conducted at a large midwestern university in 2005. The participants were recruited by e-mail notice and advertisement in newsletters on campus. The e-mail recruitment of participants went to faculty and staff through solicitations to college-level

[1] The motivation of this design is to control for any possible effects the hypothetical auction in the first round might have on the real auction in the second round. The estimation results show that the effect is negligible.

[2] The instructions are available from the authors upon request.

and university units (e.g., departments, physical plant) in order to make the recruitment pool as broadly representative of the local area and state population as possible. We restricted the pool to limit participation of graduate students and did not solicit undergraduate students. The recruitment letter indicated that participants would be asked about their market decisions on apple purchases, but nothing was said about appearance or organic production.

Seventy-four people participated in the experiment, 33% male and 67% female. The ages ranged from 20 to 70 years old, with 27% in the age 20-29 category, 30% age 30-39, 14% age 40-49, 20% age 50-59, and 9% age 60 and older. The age distribution was similar to the state average (in 2000, of the share of the state's population age 20 to 65, there were 47% in the 20-39 age range compared to the sample of 57% in this range; the state had a relatively larger share of the population in their forties). The subjects' average household income was $49,220 with a standard deviation of $30,520.[3] The median income was $42,500. This compared to the state's median household income in 1999 of $40,442. Among the participants, 17% did not have a college diploma, 11% had a college diploma, 22% had some graduate school education, and 50% had a graduate degree. The recruited sample had higher average education levels than the state average.

3. Random Effect Model

We use three sets of variables to explain the variation in WTP. First is the variation in the product quality attributes. Second is the variation in socio-demographics and consumers'

[3] Three of the observations had missing values on income, and these values were imputed using best-subset regression. The independent variables for the regression were Education, Age, Gender, and Association with the university (such as faculty, staff, student, etc.). The imputation was completed using STATA7.0.

attitudes. Third is the variation in the experiment. Based on this, we specify the following econometric model to explain the consumers' WTP for the apples:

$$WTP_{ij} = \alpha \; x_j + \beta \; y_{ij} + \gamma \; z_j + \eta_i + \varepsilon_{ij} \qquad (1)$$

where WTP_{ij} is individual *i*'s bid for product *j*; x_j is a vector of product quality attributes for product *j*, including *Organic*, *Spot*, and *OrgSpot*; *Organic* is a dummy that is one if the product is organic, and zero otherwise; *Spot* is defined as a continuous variable measuring the percentage of spot coverage;[4] *OrgSpot* measures the interaction effect between the two previous product attributes; y_{ij} is a vector of interaction effects between the socio-demographics and consumers' attitudes for individual *i* and the product quality attributes *Organic* and *Spot* for product *j*; z_j is a vector of design variables including *Picture*, *OrgPicture*, and *SpotPicture* where *Picture* is a dummy that is one for the pictures and zero for the real apples, *OrgPicture* is the interaction between *Organic* and *Picture*, and *SpotPicture* is the interaction between *Spot* and *Picture* ; and η_i is the random individual effect for the *i*th participants that captures the correlation between the bids made by the same participant. The measure η_i is assumed to follow a normal distribution with mean zero and standard deviation σ_η.

4. Results

Table 1 shows the descriptive statistics for the bids divided into two production methods (organic and conventional), four appearance levels (SpotA, SpotB, SpotC, and Spot D), and two elicitation methods (hypothetical and real auctions). There are several things that

[4] The variable *Spot* is created as a continuous variable from the four graded levels and equals the average spot level in the spot categories, i.e., *Spot* equals 0 for SpotA, 3 for SpotB, 5 for SpotC, and 9 for SpotD.

we can see directly from Table 1. First, on average, consumers are willing to pay more for organic apples than for conventional apples with the same appearance. Second, consumers on average are willing to pay more for apples with no or little cosmetic damage than for apples with more cosmetic damage. Third, consumers state higher WTP on average for all alternatives in the hypothetical auctions than in the real auction. Fourth, there are almost no zero bids for the perfect apples; in fact, none of the participants bid zero for all the apples. This indicates that the participants were willing to buy apples in the auction and that the zero bids for the spotted apples can be interpreted as zero WTP for these apples.[5] Fifth, the mean bids are below standard market prices. USDA data show the average price of fresh apples is $0.83 per pound in 1999 (Reed, Frazão, and Itskowitz, 2004), or $0.96 per pound when adjusted to 2004 apple price levels.

Table 1. Descriptive Statistics of the Bids

Production method	Auction	Statistics	SpotA	SpotB	SpotC	SpotD
Conventional	Hypothetical	Mean	2.73	2.21	1.60	0.73
		S.D.	1.60	1.26	0.98	0.74
		Median	2.25	1.90	1.38	0.59
		% zero bids	1.89%	1.89%	5.67%	37.7%
Conventional	Real	Mean	1.83	1.15	0.99	0.57
		S.D.	0.88	0.74	0.77	0.59
		Median	1.74	1.18	0.87	0.49
		% zero bids	2.06%	13.4%	18.56%	38.14%
Organic	Hypothetical	Mean	3.22	2.60	1.89	0.96
		S.D.	1.69	1.42	1.18	1.05
		Median	2.70	2.15	1.75	0.75
		% zero bids	0	0	3.77%	33.96%
Organic	Real	Mean	2.08	1.58		
		S.D.	0.95	0.88		
		Median	1.93	1.45		
		% zero bids	0	1.03%		

[5] Roosen *et al.* (1998) asked the participants to bid for an upgrade from one endowed bag of apples to other bags of apples. Thirty-five percent of the participants bid zero for all the alternatives. Their upgrade design did not allow them to distinguish between those participants who preferred the endowed bag and those who were indifferent between the bags. Bidding on all products allows us to measure both positive and negative price premiums.

The $0.96 price per pound would be $2.89 for three pounds. Our results are therefore likely to give conservative estimates of the WTP differences.

One of the initial tasks was to identify and develop measures of consumer attitudes and preferences based on the survey questions. In addition to direct responses to questions, several consumer attitudes toward quality attributes were measured as composite constructs based on the participants' degree of agreement with selected statements. The selection and ranking of the questions included in the composites were done by principal component factor analysis. To measure consumers' sensitivity to price (*Price*) we asked the participants if they agreed or did not agree with four statements about the trade-off between quality and price using a five-point Likert scale. For instance, one statement read, "I usually buy the lowest priced products." Consumers with a larger value of the index *Price* tend to be more sensitive to price of products. Other composites included consumers' concern with the environment (*Envir*), consumers' tolerance of pesticides (*Pest*), and consumers' attitude toward appearance of apples (*Appear*). Consumers with a larger value of the index *Envir* were more concerned about the environment and held stronger beliefs about the idea that organic production can improve the environment. The measure of consumers' tolerance of pesticides (*Pest*) was based on two statements concerning the safety of and restriction on pesticides. Consumers with a larger *Pest* index value were less tolerant of pesticides. The index on appearance of apples (*Appear*) is a construct based on consumers' concern about the importance of apple color, shape, texture, and size. Consumers with a larger value of *Appear* expressed more concerned about the appearance of apples. Principal component factor analysis indicated these composite constructs were uni-dimensional (all had alpha reliability of 0.6 or higher) (Cronbach, 1951).

Other measures of consumer attitudes are based on single statements. They include attitudes toward food safety (*Safe*), taste (*Taste*), and nutrition (*Nutrition*) of apples. Consumers with a larger value of each of these indexes were more concerned about the respective attributes. It is important to note that all of these measures of consumer attitudes are based on stated preferences, whereas the auctions elicit revealed preferences. The definitions of all the variables used in the segmentation are shown in Table 2.

Table 2. Definition and Summary Statistics of Variables

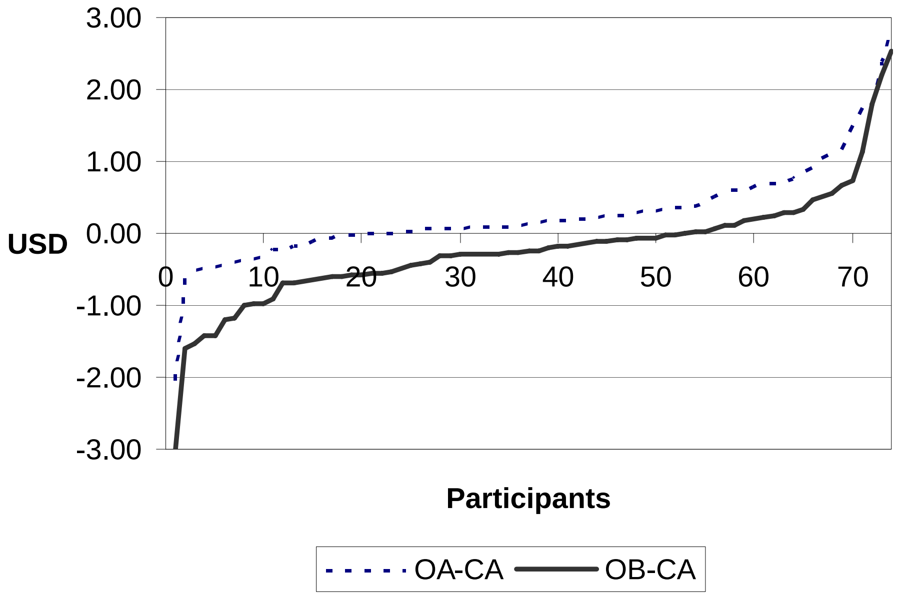
Variable	Definition	Mean	S.D.	Min	Max
Product attributes					
Organic	Organically (=1) or conventionally (=0) produced	0.46	0.50	0	1
Spot	Continuous measure of percentage coverage of spots	3.64	3.34	0	9
OrgSpot	Interaction effect of variable *Organic* and variable *Spot*				
Socio-demographics					
Age[a]	Age of the participants	40.30	13.18	25	65
Gender	Male=0, Female=1	0.67	0.47	0	1
Edu[b]	Education on a 6-point scale	4.96	1.27	2	6
Income	Income in thousands of dollars	49.22	30.52	7.5	120
Attitudes					
Price[c]	Price sensitivity	0	1	-2.37	1.98
Envir[c]	Concern about environment	0	1	-2.29	1.89
Pest[c]	Pesticides risk tolerance	0	1	-2.41	2.33
Appear[c]	Attitude towards appearance of apples	0	1	-2.64	1.84
Taste[d]	Taste of apples	4.70	0.58	3	5
Safe[d]	Food safety	4.05	1.13	1	5
Nutrition[d]	Nutrition of apples	3.48	1.22	1	5

[a] The age variable has seven categories and we have used the midpoint of the categories to form a continuous variable.
[b] 1 = Some high school, 2 = High school diploma, 3 = Some college or less, 4 = College diploma, 5= Some graduate school, and 6 = Graduate degree.
[c] Factors
[d] Based on the answer to the following question: How important are the following attributes of apples when you decide which apples to buy? (5-point scale where 1 is not important and 5 is very important)

Figure 1 shows the cumulative distribution of the difference in WTP between Organic A and Conventional A (OA-CA), and Organic B and Conventional A (OB-CA). OA-CA is calculated by subtracting the individual participant's mean bid for Conventional A from the same participant's mean bid on Organic A. Similarly, OB-CA is calculated by subtracting the individual participant's mean bid for Conventional A from the same participant's mean bid on Organic B.

From Figure 1, we can see that 19 (25%) of the participants bid higher for the Conventional A apples than the organic apples with the same appearance. This indicates that these consumers think that there is a negative value associated with organic production. Of the 55 (75%) participants bidding more for the Organic A than the Conventional A, 18 (24%) bid more than 50¢ more, eight (11%) bid more than $1 more, and two (3%) bid more than $2 more for the organic apples.

Figure 1. Cumulative Distribution of the Difference in Willingness to Pay between Organic A and B and Conventional A



Comparing Organic B with Conventional A, we can see that of the 55 participants preferring the organic apples when they had the same appearance, now only 21 (28% of the total sample) still prefer the organic apples. This drastic decline in the group preferring the organic apples indicates that the appearance is very important for many consumers.

Table 3 includes information about the socio-demographic and attitude variables across the three consumer groups indicated by Figure 1. Group 1 prefers conventional to organic (Bid Conventional A > Bid Organic A), group 2 prefers organic but only if the appearance is as good as for the conventional (Bid Organic A > Bid Conventional A > Bid Organic B), and group 3 prefers the organic even when the appearance is lower than that of the conventional (Bid Organic B> Bid Conventional A). We can see that the participants in group 1 tend to be younger than those in other groups and they are less concerned about the food safety–related attributes such as environment, pesticides, food safety and so on. And they are the group that has the lowest income level and is most concerned about price; that is, they are the group with the highest sensitivity to price. The consumers in group 2 care more about appearance than do the other groups. They are almost neutral to environment and pesticides. In contrast, those in group 3 value the food safety–related attributes and taste the most, and they value appearance and price the least. Those in group 3 have the highest income and education levels, and they are the oldest compared with the other two groups.

To see if these groups differ significantly in the socio-demographics and attitudes, MANOVA and the Wilk's Λ^* test are used (Johnson and Wichern, 2002). ANOVA is employed to test if the groups differ in each of the variables. The P-values of the tests are

Table 3. Socio-demographics and Attitudes of the Groups

	Group 1		Group 2		Group 3		ANOVA	MANOVA
	Mean	S.D.	Mean	S.D.	Mean	S.D.	P-value	P-value
Socio-demographics								
Age	36.05	15.23	38.23	12.24	43.10	12.49	0.21	
Gender	0.63	0.49	0.71	0.46	0.67	0.48	0.84	
Edu	4.84	1.30	5.00	1.15	5.10	1.44	0.82	
Income	36.97	22.75	41.14	31.06	58.69	32.85	0.05	0.51[b]
Attitudes								
Price[a]	0.59	1.19	-0.12	0.70	-0.34	1.08	0.01	
Envir[a]	-0.52	1.01	0.02	1.01	0.46	0.79	0.01	
Pest[a]	-0.22	1.12	-0.05	0.96	0.28	0.97	0.27	
Appear[a]	-0.17	1.10	0.27	1.01	-0.31	-0.84	0.08	
Taste[a]	-0.34	0.98	-0.01	0.94	0.37	1.06	0.08	
Safe[a]	-0.16	0.98	0.10	1.03	0.32	0.98	0.22	
Nutrition[a]	0.12	0.96	-0.18	1.07	0.02	0.94	0.35	0.04[c]
								0.20[d]

[a] These attitudes variables are standardized.
[b] MANOVA of socio-demographic variables.
[c] MANOVA of attitude variables.
[d] MANOVA of both socio-demographic and attitude variables.

listed in Table 3. The Wilk's Λ^* test statistic when including all the variables is 0.66, and the corresponding P-value is 0.20. So the null hypothesis that the mean vectors are the same across the groups cannot be rejected at the 5% significance level. The P-value of the Wilk's Λ^* test statistic obtained by including only the socio-demographic variables is 0.51, so the null hypothesis that the groups are the same in socio-demographic variables cannot be rejected at the 5% significance level. However, the ANOVA results show that the three groups differ in income at the 5% significance level. The P-value of the Wilk's Λ^* test statistic when including only the attitude variables is 0.04, so the null hypothesis that the groups are the same in their attitudes toward food safety–related quality attributes, price, appearance, taste, and nutrition is rejected at the 5% significance level. We find that the three groups differ in attitudes at the 5% level of significance. Thus, from the ANOVA results we conclude that the three groups differ mainly in their attitudes toward price, environment, appearance, and taste, and by their income levels.

It is useful to compare our results with those of Roosen *et al*. (1998); the two studies were done 10 years apart and both were conducted in the state of Iowa. Although the two studies differ in many aspects, they address a similar valuation problem. In the Roosen *et al.* study, 38% of the participants had a high degree of concern about pesticide use, and of these, 76% preferred stricter pesticides regulations. In our study, 42% of the participants were (very) concerned about pesticide use; of these, 88% think stricter pesticide regulations should be set. Furthermore, Roosen *et al.* find that 35% of participants consistently bid zero for all potential upgrades from conventional apples to apples produced with no pesticides. In contrast, in our study, 25% of the participants bid less for the Organic A than for the Conventional A.

4.1 The Random Effect Models

We estimated three random effects models. All three models include the product attribute and experimental design variables. In addition, Model 1 includes socio-demographic interaction effects, Model 2 includes attitude interaction effects, and Model 3 includes both socio-demographic and attitude interaction effects. We estimated the three models to check the robustness of the estimation and to avoid any identification problems.[6] The models include only the interaction effects of the socio-demographic and/or attitude

[6] We estimated the correlation between the socio-demographic variables and preference attitude variables and found that the largest correlation was 0.17. We conclude that there is no multicollinearity problem in Model 3. However, since we have only 74 participants and Model 3 includes 28 independent variables, there may be identification problems. For this reason we include both Model 1 and Model 2 in the analysis.

variables with *Organic* and *Spot*; the effects of the socio-demographic and/or attitude variables alone have been largely captured by the individual random effect.[7] The results from the three models are quite similar, which indicates that the estimates are robust toward small changes in the model specification.

Maximum likelihood was used to estimate the parameters in equation (1). To simplify the interpretation of the parameters associated with the quality attributes, the variables that interact with them are standardized with a mean of zero and a standard deviation of one. The standardization is done by subtracting the respective variable's mean and dividing by its standard deviation. The estimated parameters from the random effects models are shown in Table 4.

From Table 4 we can see that the apple quality attributes (organic and spot) affect the consumers' WTP for apples and the results are statistically significant. The three models get almost identical results for the main effects. Also, the constant, which can be interpreted as the average bid for 3 pounds of conventional apples without any spots, is very similar among the three models: about $1.74. The production method affects consumer WTP significantly. Consumers are willing to pay more for organic apples than for conventional apples: the premium for organic apples without any spots is about $0.35 per 3 pounds ($0.12/pound). However, the interaction between organic production and cosmetic damage (level of spots) is statistically significant: the premium for organic production decreases $0.04 per 3-pound bag when the level of spot damage increases by 1%. Taking account of the combined direct and indirect effects, the consumer WTP

[7] We tried another model that included both the individual socio-demographic variables and attitude preference variables and the interaction effects. To test the model specification, a likelihood ratio test was conducted. The test statistic was 6, which is less than the critical value 19.68, so the null hypothesis that the coefficients of the individual socio-demographic variables and attitudes preference variables are zero cannot be rejected at the 0.05 significance level.

Table 4. WTP for (Organic) Apples with Spots, Random Individual Effect Models[a]

	Model 1			Model 2			Model 3		
Variables	**Coef.**	**S.D.**	**P-value**	**Coef.**	**S.D.**	**P-value**	**Coef.**	**S.D.**	**P-value**
Product attributes									
Constant	1.74***[b]	0.10	0.00	1.75***	0.09	0.00	1.74***	0.09	0.00
Organic	0.35***	0.08	0.00	0.34***	0.08	0.00	0.35***	0.08	0.00
Spot	-0.14***	0.01	0.00	-0.14***	0.01	0.00	-0.14***	0.01	0.00
OrgSpot	-0.04***	0.01	0.00	-0.04***	0.01	0.00	-0.04***	0.01	0.00
Socio-demographic interaction effects									
EduOrg[c]	-0.04	0.05	0.48				-0.04	0.05	0.46
GenderOrg	0.001	0.05	0.98				-0.05	0.05	0.32
AgeOrg	-0.07	0.06	0.19				-0.07	0.07	0.28
IncomeOrg	0.20***	0.06	0.00				0.14**	0.07	0.03
AgeSpot	-0.01	0.01	0.26				-0.01	0.01	0.24
GenderSpot	-0.02**	0.01	0.02				-0.02*	0.01	0.02
EduSpot	0.01**	0.01	0.08				0.01	0.01	0.19
IncomeSpot	0.003	0.01	0.73				0.001	0.01	0.87
Attitude interaction effects									
PriceOrg				-0.11**	0.05	0.05	-0.09*	0.06	0.10
EnvirOrg				0.20***	0.07	0.01	0.21***	0.07	0.01
SafeOrg				0.25***	0.07	0.00	0.22***	0.07	0.00
PestOrg				-0.009	0.07	0.89	-0.02	0.07	0.78
AppearOrg				0.01	0.05	0.82	0.02	0.05	0.68
TasteOrg				0.26***	0.07	0.00	0.24***	0.07	0.00
NutritionOrg				0.009	0.05	0.86	-0.005	0.05	0.92
PriceSpot				-0.01	0.01	0.19	-0.01	0.01	0.37
EnvirSpot				-0.006	0.01	0.59	-0.002	0.01	0.89
SafeSpot				0.001	0.01	0.91	0.005	0.01	0.63
PestSpot				0.008	0.01	0.43	-0.001	0.01	0.88
AppearSpot				-0.03***	0.01	0.00	-0.02***	0.01	0.00
TasteSpot				-0.001	0.01	0.92	0.005	0.01	0.64
NutritionSpot				-0.001	0.01	0.89	0.002	0.01	0.82
Experimental design effects									
Picture	1.04***	0.10	0.00	1.01***	0.10	0.00	1.04***	0.09	0.00
PictureOrg	0.14	0.11	0.18	0.18*	0.11	0.08	0.14	0.10	0.18
PictureSpot	-0.08***	0.02	0.00	-0.08***	0.02	0.00	-0.08***	0.02	0.00
Model parameters									
σ_η	0.61***	0.06	0.00	0.59***	0.05	0.00	0.60***	0.06	0.00
σ_ε	0.80***	0.02	0.00	0.78***	0.02	0.00	0.77***	0.02	0.00
ρ	0.37	0.04		0.37	0.04		0.37	0.05	

Note: The models are random individual effect models estimated with STATA 7.0.

[a] Likelihood ratio test statistics for the goodness of fit of the models are 619 for Model 1, 662 for Model 2, and 679 for Model 3, and the p-values are less than 0.0001.

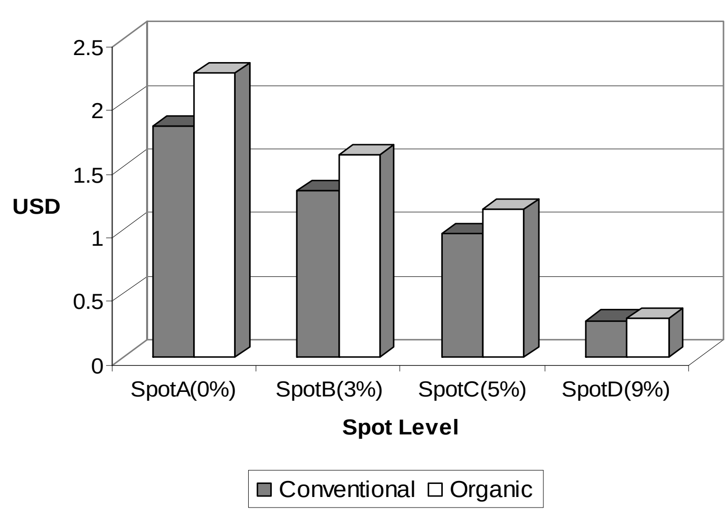
[b] Single, double, and triple asterisks (*) denote significance at the 0.1, 0.05, and 0.01 levels, respectively.

[c] *EduOrg* means the interaction effect between variable *Edu* and *Organic* and *EduSpot* means the interaction effect between variable *Edu* and *Spot.* Similar definitions hold for the attitude interaction effect variables and experiment design variables as well.

decreases by $0.14 per 3-pound bag when the level of spot damage for conventional apples increases by 1%. For organic apples, when the level of spot damage increases by 1%, the consumer WTP decreases by $0.18. The difference in the discount between the two production methods is statistically significant.

Figure 2 summarizes the consumer WTP for 3 pounds of organic apples and conventional apples with different levels of spots used in our experiment. Note that consumers' WTP for Organic B apples is less than that for Conventional A apples; consumers' WTP for Organic D apples is less than that for Conventional C apples. We can conclude that consumers make a trade-off between production method and the blemish level of the apples. Even though, in general, consumers are willing to pay more for organic apples, when there are "too many" blemishes on the organic apples, consumers prefer to buy conventional apples. An extrapolation of the numbers shown in Figure 2 to apples with even more spots than the amounts on Spot D apples (9%) shows that consumers would not be willing to pay for such apples regardless of the production method.

Figure 2. Consumer WTP for Apples with Different Levels of Spots



As shown in Table 4, the interactions between the socio-demographic variables or attitude variables and the production methods or damage levels show some statistically significant interaction effects. Model 1 and Model 3 indicate that the interaction effect between income and organic production methods is positive and significant. Those who have higher income are willing to pay a higher price premium than those who have a lower income level. Other interactions with the production method are not statistically significant.

Model 1 and Model 3 show that two of the socio-demographic interactions with the spot damage are statistically significant. The interaction effect between gender and spot damage is negative and significant at the 5% significance level; females are more reluctant to buy apples with spots. One possible explanation for this might be that females show more concern about the aesthetics of food than do men. Or, perhaps more time and experience in grocery shopping on average (Hamrick and Shelley, 2005) contribute to females' lower bids for apples with more spots. The significant interaction effect between education level and spot level in Model 1 indicates that those with higher education levels are more willing to buy apples with spots.

Model 2 and Model 3 show that several of the interaction effects between organic production methods and attitude variables are significant. Those who are less sensitive to price (*PriceOrg*) and consumers with greater concerns about the environment (*EnvirOrg*) are willing to pay a higher price premium for organic products compared to others. Positive interaction with environmental concerns suggests positive association between organic production methods and environmental interests in the minds of consumers. In addition, consumers' concern about food safety is also positively associated with organic production (*SafeOrg*). Those who are more concerned about food safety are willing to pay a higher premium for organic apples than those who are less concerned with food safety, a result that suggests that consumers think organic products are safer than conventional products. Finally, those who are more concerned with taste (*TasteOrg*) are willing to pay a higher premium for organic product. In summary, willingness to pay more for organic products is enhanced by consumers' being less sensitive to price, more concerned with the environment, more concerned with the safety of food products, and their having high levels of interest in the "tastiness" of food products.

These results are consistent with previous studies showing that consumers associate organic production methods with a reduced health risk and may chose to reduce the risk from pesticide residues by switching to organically grown products (Williams and Hammitt, 2001; Magnusson and Cranfield, 2005). Recent survey evidence shows that consumers purchase organic foods because they perceive the foods to be fresh (68.3%), better for health, and a better source of nutrition (67.1%) (Whole Foods Market, 2005). Over 70% (70.3%) of those surveyed said they bought organic food or beverages in order to avoid pesticides.

The interaction effects between spot level and attitude constructs are less strong, though similar between Models 2 and 3. The interaction effect between concerns about appearance and spot damage (*AppearSpot*) is negative and statistically significant for Model 2 and Model 3. Those who are more concerned with appearance place a higher discount for apples with increased levels of spot damage.

Controlling for the experimental design was important. The variable *Picture* is highly significant in all three models. Relative to the average bid of $1.74 for a 3-pound bag of conventional apples without spots, participants bid about one dollar more for apples presented in pictures than for real apples. In this respect, our results are in line with the large literature on hypothetical bias in valuation studies. Because the participants did not need to pay the price they bid to buy the product when presented with pictures, they tended to overbid for pictures compared to the cases where they were presented with real products and faced the chance they would need to pay out of pocket for the real product.[8] The interaction effect between *Organic* and *Picture* (*PictureOrg*) is positive and significant in Model 2 and the interaction effect between *Spot* and *Picture* (*PictureSpot*) is negative and significant in all three models, which indicates that the hypothetical bias is not fixed with the changes in WTP. Actually, from Table 1 we get similar results. Hypothetical bias seems to be proportional to WTP. If we divide the sum of the mean WTP for apples shown in a picture by the sum of the mean WTP for real apples, we find that the WTP in the hypothetical auction is 1.6 times that of the real auction.

Both $\hat{\sigma}_\eta$ and $\hat{\sigma}_\varepsilon$ are significant. A likelihood ratio test was conducted concerning the individual random effect (null hypothesis: $\sigma_\eta=0$). The test statistic had values of 290 for Model 1, 295 for Model 2, and 299 for Model 3. All of the corresponding P-values were < 0.001 and the null hypothesis was rejected for each model. That is, the individual random effects cannot be ignored and need to be included in order to estimate the results accurately. The $\hat{\rho}$ value for the models was 0.37. This parameter measures the correlation between the bids on different apples by the same participants.

[8] We ran another model using the bids on pictures of apples only and found the constant (the average bid for 3-pound conventional perfect apples) was $2.78 instead of $1.74, and the premium for organic was $0.46 instead of $0.35 for 3 pounds of apples.

5. Conclusions

Consumers want environmental friendly production methods, but they do not want the natural consequences of the environmental friendly production: the blemished appearance of products. This result is of course very troublesome for organic producers. Organic producers are less able to avoid problems with cosmetic appearance and they are hit harder in the retail market if they produce less-than-perfect apples. At first this result is somewhat surprising, given that previous studies have shown that the majority of consumers say they buy organic products to avoid pesticides. However, since the consumers are willing to pay more for perfect organic apples than for perfect conventional apples, a percentage discount from cosmetic damage yields a higher dollar value in the discount of organic apples than for the conventional apples.

Of specific interest in this study is the premium that consumers are willing to pay for organic apples and the effect of different levels of cosmetic damage on the premium. We find that the premium for organic apples decreases as the level of spots increases, a result that supports earlier findings of Thompson and Kidwell (1998) and Roosen *et al.* (1998). Furthermore, our experimental design allows us to estimate interaction effects between production method and cosmetic damage. We find not only that the negative effect from cosmetic damage offsets the positive effect from organic production but also that cosmetic damage leads to discounting the premium for organic production. Consumers' tolerance of cosmetic damage on apples is limited. Even at relatively low levels of blemishes on the surface of organic apples, consumers preferred perfect-looking conventional apples. The consumers differ with respect to how they rank the importance of appearance. There is a relatively large segment of consumers in the organic market who are willing to accept a small level of cosmetic damage. However, if apple growers try to sell less-than-perfect organic apples at a price that is significantly above the going price of conventional apples, very few consumers will be willing to buy the organic apples.

This finding suggests the importance of quality attributes connected to cosmetic appearance, as is the case today with the fruit grading system of U.S. Department of Agriculture, Agricultural Marketing Service, and that exists in many private contracts for produce. To a large extent, fresh fruits in U.S. grocery stores have uniform appearance, while the fruits with imperfect appearance often are diverted as processed product such as fruit juice and sauce. Our findings show that even when there is no strict federal grading system, fresh fruits with cosmetic damage have little potential in today’s retail market because of consumers’ limited tolerance for imperfect cosmetic attributes. When faced with limited consumer tolerance for cosmetic damage, apple producers must account for the trade-off between production technology and cosmetic damage in their production decisions in order to ensure their profits.

References

Acebron, L.B., and D.C. Dopico. 2000. “The Importance of Intrinsic and Extrinsic Cues to Expected and Experienced Quality: An Empirical Application for Beef.” *Food Quality and Preference* 11:229–38.

Alfnes, F., and K. Rickertsen. 2003. “European Consumers’ Willingness to Pay for U.S. Beef in Experimental Auction Markets.” *American Journal of Agricultural Economics* 85(2): 396–405.

Alfnes, F., A. Guttormsen, G. Steine, and K. Kolstad. 2006. “Consumers’ Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives.” *American Journal of Agricultural Economics* 88(4):1050-1061.

Baker, G. 1999. “Consumer Preferences for Food Safety Attributes in Fresh Apples: Market Segments, Consumer Characteristics and Marketing Opportunities.” *Journal of Agricultural and Resource Economics* 24(1): 80-97.

Blend, J.R., and E.O. van Ravenswaay. 1999. “Measuring Consumers’ Demand for Ecolabeled Apples.” *American Journal of Agricultural Economics* 81(5): 1072–1077.

Darby, M.R., and E. Karni. 1973. “Free Competition and the Optimal Amount of Fraud.” *Journal of Law and Economics* 16:67-88.

Hamrick, K., and K.J. Shelley. 2005. “Data Feature: How Much Time Do Americans Spend Preparing and Eating Food?” *Amber Waves*, magazine of the Economic Research Service, U.S. Department of Agriculture, pp. 10-11. Available at http://www.ers.usda.gov/AmberWaves/November05/DataFeature/ (accessed November 2006).

Johnson, R.A., and D.W. Wichern 2002. *Applied Multivariate Statistical Analysis*, 5th ed. Englewood Cliffs, NJ: Prentice-Hall International.

Larue, B., G.E. West, C. Gendron, and R. Lambert. 2004. “Consumer Response to Functional Foods Produced by Conventional, Organic, or Genetic Manipulation.” *Agribusiness* 20(2):155-166.

Loureiro, M.L., J.J. McCluskey, and R.C. Mittlehammer. 2001. “Assessing Consumer Preferences for Organic, Eco-Labeled, and Regular Apples.” *Journal of Agricultural and Resource Economics* 26:404–16.

Lusk, J.L., T. Feldkamp, and T.C. Schroeder. 2004. “Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods.” *American Journal of Agricultural Economics* 86: 389–405.

Lusk, J.L., L.O. House, C. Valli, S.R. Jaeger, M. Moore, B. Morrow, and W.B. Traill. 2004. “Effect of Information about Benefits of Biotechnology on Consumer Acceptance of Genetically Modified Food: Evidence from Experimental Auctions in United States, England, and France.” *European Review of Agricultural Economics* 31:179–204.

Magnusson, E., and J.A.L. Cranfield. 2005. “Consumer Demand for Pesticide Free Food Products in Canada: A Probit Analysis.” *Canadian Journal of Agricultural Economics* 53(1, March): 67-81.

Melton, B.E., W.E. Huffman, J.F. Shogren, and J.A. Fox. 1996. “Consumer Preferences for Fresh Food Items with Multiple Quality Attributes: Evidence from an Experimental Auction of Pork Chops.” *American Journal of Agricultural Economics* 78:916-923.

Reed, J., E. Frazão, and R. Itskowitz. 2004. "How Much Do Americans Pay for Fruits and Vegetables?" Agriculture Information Bulletin No. AIB790, July. Electronic Report from the Economic Research Service, USDA. Available at http://www.ers.usda.gov/publications/aib790/ (accessed November 2006).

Roosen, J., J.A. Fox, D.A. Hennessy, and A. Schreiber. 1998. "Consumers' Valuation of Insecticide Use Restrictions: An Application to Apples." *Journal of Agricultural and Resource Economics* 23: 367-384.

Rozan, A., A. Stenger, and M. Willinger. 2004. "Willingness-to-Pay for Food Safety: An Experimental Investigation of Quality Certification on Bidding Behaviour." *European Review of Agricultural Economics* 31(4):409-425.

Thompson, G.D., and J. Kidwell. 1998. "Explaining the Choice of Organic Produce: Cosmetic Defects, Prices, and Consumer Preferences." *American Journal of Agricultural Economics* 80: 277-287.

Umberger, W.J., and D.M. Feuz. 2004. "The Usefulness of Experimental Auctions in Determining Consumers' Willingness to Pay for Quality Differentiated Products." *Review of Agricultural Economics* 26: 1–16.

Waugh, F.V. 1928. "Quality Factors Influencing Vegetable Prices." *Journal of Farm Economics* 10(2, April): 185-196.

Wei, S., S. Singgih, E.J.Woods, and D. Adar, 2003. "How Important Is Appearance? Consumer Preferences for Mandarins in Indonesia." *International Journal of Consumer Studies* 27:406-411

Whole Foods Market. 2005. “Nearly Two-Thirds of Americans Have Tried Organic Foods and Beverages.” Press release, November 18. http://www.wholefoodsmarket.com/company/pr_11-18-05.html

Williams, P.R.D., and J.K. Hammitt. 2001. “Perceived Risks of Conventional and Organic Produce: Pesticides, Pathogens, and Natural Toxins.” *Risk Analysis* 21(2, April): 319-330.