In-Store Valuation of Steak Tenderness

Jayson L. Lusk, John A. Fox, Ted C. Schroeder, James Mintert, and Mohammad Koohmaraie

Experimental methods were used to examine consumer willingness-to-pay for steak tenderness in a grocery store setting. When relying on a taste test alone to determine product quality, the participants paid an average premium of $1.23/lb for a tender versus tough steak. Fifty-one percent of the participants were willing to pay an average of $1.84/lb when they had completed a taste test and were also provided information about the steak's tenderness. Results indicate that most consumers prefer more tender steaks and that many are willing to pay a premium for tender steaks.

Key words beef, consumer demand, experimental economics, field experiment, tenderness, willingness to pay

The beef sector has faced weakening demand for two decades, with demand declining every year from 1979 to 1998 (Purcell). U.S. Department of Agriculture (USDA) data indicate per capita beef consumption dropped 20% from 1970 to 1998 while inflation-adjusted retail beef price declined 25% (LMIC). From 1990 to 1998 alone beef demand declined by 18%.[1] The dramatic demand decline has contributed to considerable restructuring and downsizing of the beef industry. For example, from January 1975 (the U.S. beef cow inventory peak) to January 1999, the U.S. beef cow herd declined by 27% (LMIC).[2]

Numerous studies have provided insights into possible causes of the precipitous decline in beef demand (e.g., Brester and Schroeder; Brester and Wohlgenant; Capps, Moen, and Branson; Chavas; Eales and Unnevehr 1988 and 1993; Flake and Patterson; Kinnucan et al.; Lamb and Beshear; Moon and Ward; Purcell; Unnevehr and Bard). Factors cited by these authors as contributing to the demand decline include changes in relative prices, consumer health concerns, food safety concerns, product convenience and offering, product quality and consistency, changing demographics, and evolving consumer preferences. Notable among these factors are product quality issues. Although changes in consumer "tastes and preferences" are exogenous to the beef industry, product quality and consistency are attributes that can be controlled during production and processing, albeit at some cost. Improvements in beef quality and consistency, however, have been slow.

Surveys of beef packers, purveyors, restauranteurs, and retailers indicate product uniformity, consistency, and tenderness are among the highest ranked beef quality concerns (Smith et al., 1995). Among these quality traits, tenderness has been demonstrated as the most important palatability attribute of beef (Dikeman; Huffman et al.; Miller et al.). However, current USDA quality grading standards are ineffective at identifying meat tenderness (Savell et al.). The current grading system uses intramuscular fat or marbling as a measure of quality. Marbling, however, is poorly correlated with tenderness. Wheeler, Cundiff, and Koch concluded that marbling explained at most only 5% of the variation in beef tenderness. This inconsistency in quality identification has led many industry participants to the same conclusions as Wheeler, Cundiff, and Koch

The authors are, respectively, assistant professor, Mississippi State University, associate professor, professor, and professor, Kansas State University and research leader, Meat Research Unit, USDA-ARS U.S. Meat Animal Research Center.

The authors acknowledge the Research Institute on Livestock Pricing for funding. The authors thank Jason Shogren and two anonymous reviewers for their helpful comments and insights. The authors are grateful to Jenny Graff, Matt Hoobler, and Christy Lusk for assistance with data collection. Mention of a trade name, proprietary product or specific equipment is necessary to report factually on the available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by the USDA implies no approval of the product to the exclusion of others that may also be suitable.

[1] Demand decline estimate is the percentage decline in real retail price (from 1990 to 1998) holding per capita quantity constant at the 1990 level and assuming an own-price demand elasticity of −0.67.

[2] Although the beef cow herd has declined over this time period, beef production has changed little due to technological change, increases in feeding efficiency, and lower feed prices.

Copyright 2001 American Agricultural Economics Association

(p. 3150) that, "USDA quality grade does not sufficiently segregate carcasses for palatability differences, and thus a direct measure of meat tenderness is needed to supplement USDA quality grade."

Recent improvements in technology have enabled researchers to more effectively segregate carcasses into tenderness categories, thus providing a means of increasing product quality and consistency to the consumer. Additionally, more appropriate price signals may be passed from retail to farm levels by improving quality measurement. Improvements of this nature provide the potential to counteract declining demand by helping the beef industry provide consumers a product that possesses the quality attributes they desire.

New beef grading systems will only be based on meat tenderness if consumers value this attribute. In this study, we use an experiment to estimate consumer willingness-to-pay for a higher level of steak tenderness, and estimate the influence of economic and demographic factors on willingness-to-pay values.[3] We also show how revealing information about tenderness affects both preference and willingness-to-pay for tender steaks.

Data were obtained using a novel experimental method with consumers shopping in retail grocery stores. Following a taste test, steaks of varying tenderness were auctioned using a variant on the true-value revealing auction mechanism proposed by Becker, DeGroot, and Marschak. Results indicate that: (a) consumers readily distinguish between tenderness levels in a blind taste test, (b) many, but not all, consumers are willing to pay a premium for more tender relative to less tender steaks, and (c) when information regarding tenderness is revealed to consumers in addition to taste tests, they are significantly more likely to prefer and pay more for tender steak. These results have important implications for beef producers and processors, especially firms considering investing in tenderness measurement technology to market steaks identified by tenderness levels.

A Tenderness Based Grading System

One direct method for measuring tenderness is the Warner–Bratzler (WB) shear force test. The WB test measures the amount of force required to penetrate a cooked cut of meat and assigns a numerical value to a cut of steak indicating its tenderness level. The WB method explains more of the variation in meat tenderness than any other tenderness testing system (Shackelford, Wheeler, and Koohmaraie). Recently, Koohmaraie et al. developed a beef processing system that incorporates the WB test and can be used in a commercial processing plant. Based on findings from professional taste panels, Koohmaraie et al. proposed that the tenderness grading system be used to segregate carcasses into three tenderness classifications–guaranteed tender, intermediate tender, and probably tough.[4] The system segregates carcasses with 90% accuracy, higher than any system previously devised (Shackelford, Wheeler, and Koohmaraie). In addition to providing a more consistent prediction of eating quality, this tenderness classification system results in more carcasses being segregated into the highest quality grade than under current USDA quality grading standards. This occurs because some carcasses that are graded as Choice or Select in the current grading system would be categorized as guaranteed tender under this alternative grading system.

Methods

We used an experimental market procedure to elicit consumer willingness-to-pay values for a higher level of tenderness in beef steaks. Valuation in non-hypothetical experiments has unique advantages (Fox et al.) and the experimental valuation of food attributes is well established in the literature (for examples see Shogren et al., 1994; Hayes et al.; Melton et al.; Roosen et al.). To date, most, if not all, such valuation experiments have been conducted in laboratory settings with groups of participants randomly selected

[3] In the only other study to report on consumer valuation of tenderness, Boleman et al provided steaks of varying tenderness to 42 families. Participants readily distinguished between the tough and tender steaks In a second phase of that study, 19 purchased steaks. A majority (55%) of the steaks purchased were in the most-tender category but the study did not report the number of families purchasing different steak types

[4] One important change that would occur with such a tenderness-based grading system is that a larger percentage of beef carcasses would receive the highest quality grade than under the current system. Shackelford et al predict that 29% of grain-fed carcasses would grade guaranteed tender while the current system grades carcasses as 1% Prime, 47% Choice, 47% Select, and 5% Standard.

from the general population. Values are typically elicited using a Vickrey second price auction where participants submit a sealed bid, and the auction winner pays the second highest bid price (Vickrey). In this study, the valuation setting is moved to more familiar territory for the consumer by conducting the experiment in a grocery store with individual shoppers rather than in a lab with randomly selected groups of participants.

In-store valuation has demonstrable and potential advantages for the experimenter compared to a lab setting. First, in selecting subjects, it allows one to more closely target the population of interest, in this case meat buyers, by conducting the experiment at the meat case where the relevant purchase decisions are normally made. Sample selection bias may still arise since not every shopper will participate, but that bias will likely be smaller than in laboratory experiments because participation involves less inconvenience for the subject. Compensating for participant inconvenience in laboratory sessions requires financial inducements, typically around $20 to $30. Recruitment without compensation likely increases selection bias, but, since opportunity costs vary across individuals, it is possible that uniform compensation may differentially impact subjects' revealed values. Buzby et al. (1998) reported a significant positive effect on revealed values for reductions in *Salmonella* risk when a $3 participation payment was made to student subjects whose opportunity costs were likely near 0.

A clear advantage to the in-store experiment is the reduced cost when $20–$30 participation payments are eliminated. Although some form of inducement for subjects to participate is still required, the total cost per observation is likely to be much lower. In this study the inducement we provided was integral to the valuation exercise. Therefore, recruitment and participation expenses were effectively zero. Facilitating larger sample sizes for a given research budget addresses an issue, i.e., small samples, that has long been a concern in economic experiments.

The switch to valuation in an individual rather than a group setting also required an alternative valuation mechanism since the Vickrey auction only works for a group. We elicited individual values using a variation on the Becker–DeGroot–Marschak (BDM) mechanism (Becker, DeGroot, and Marshak), which, like the Vickrey auction, provides incentives for subjects to accurately represent their preferences. The incentive structures in the Vickrey and BDM mechanisms are similar in that bids are separated from market price. In the Vickrey auction the high bidder pays a price equal to the second highest bid; in BDM, participants receive the item if their bid exceeds a subsequently drawn random number, where the randomly drawn number becomes the market price. The BDM procedure is commonly used to elicit certainty equivalents for lotteries in risk preference studies (e.g., Kachelmeier and Shehata) and has also been used to value private goods (Boyce et al.).[5]

To facilitate our experiment in the retail store, we modified BDM by using a predetermined price for the good being auctioned, i.e., an upgrade from a tough steak (provided free for participating in the experiment) to a tender steak. This predetermined price was unknown to the subjects and thus independent of their stated bid. The price was revealed as the purchase price only to subjects whose bid exceeded the price.[6] When the price is higher than a participant's bid, they are pleased not to pay for the upgrade; when the price is lower, they are pleased to obtain the upgrade for an amount below their maximum willingness to pay.

Procedures

Data were collected from shoppers at three urban retail grocery stores, owned by a large regional chain, in the midwestern United States. The experiment involved five steps in each of two treatments. Consumers participated one at a time without any knowledge of the previous respondent's outcomes. After completing the experiment, participants were asked not to discuss the results of the experiment with other shoppers in the store.

Step 1. Shoppers approaching the meat counter were asked to participate in a short

[5] Recent work by Shogren et al (2000) emphasizes the provision of a market environment with feedback in valuation studies There seems to be a trade-off here between (a) the ability to provide feedback from the market, and (b) the ability to create an auction environment in-store that is relatively quick and not totally disruptive to the store environment We feared that using repeated trials with market feedback would make the experiment a spectacle (or circus) that might lead bidders to feel like they were on display and consequently affect bidding behavior in uncontrollable ways

[6] Price for the upgrade was typically between $0 50 and $1 00

experiment for which they would receive a free 12 oz ribeye steak.

Step 2. Participants completed a short written survey that required disclosure of basic demographic information including age, gender, household size, household income, education level, and preference for steak doneness and USDA quality grade.

Step 3. Participants then sampled two different types of steaks labelled Red or Blue: Red was "Guaranteed tender" (based on the slice shear force test) and Blue was "Probably tough."[7] In experimental treatment 1, consumers were not told that the samples differed in tenderness—they had to make this assessment independently.

Step 4. Participants next responded to questions about which steak they preferred overall and for the individual attributes of taste, tenderness, texture, and juiciness.

Step 5. Participants were given, free of charge, a 12 oz. Blue (probably tough) ribeye steak. If they preferred the Blue (probably tough) steak, the experiment ended. If they indicated a preference for the Red (guaranteed tender) steak, they were asked (both verbally and on the survey) to indicate the *most* they would be willing-to-pay to exchange their Blue (probably tough) steak for the 12 oz. Red (guaranteed tender) ribeye steak. Respondents were told that if their bid exceeded a predetermined price (unknown and exogenous to them), they would make the exchange at that predetermined price. If their bid was less than the predetermined price, then they kept their Blue (probably tough) steak.

The second treatment was identical to the first except that the words Red and Blue were replaced with "guaranteed tender" and "probably tough", respectively. That is, the consumers were provided information about the steaks in addition to their taste sampling. Koohmaraie et al. used the descriptions "guaranteed tender" and "probably tough" when segregating steaks into different tenderness categories. To provide participants with a consistent explanation of the tenderness classification system, we provided the following statement:

> The USDA has developed a technology to categorize steaks according to tenderness. The classification system uses slice shear force to give an actual value of steak tenderness. Steaks are separated into different categories according to shear force values. The three categories are: Guaranteed Tender, Intermediate Tender, and Probably Tough.

The steaks used for sampling were deemed tender or tough according to a slice shear test using the procedures outlined in Shackelford, Wheeler, and Koohmaraie. Meat scientists at the USDA's Meat Animal Research Center at Clay Center, Nebraska conducted tenderness measurement and categorization of the steaks.

A flow diagram of the experiment is shown in figure 1. Consumer choices in the experiment are based on their ability to "taste" tenderness, their preference for tenderness, their attitudes toward beef in general, and previous experience with beef. These factors may be captured by various consumer demographics, consumption habits, and economic characteristics. Factors such as age, education, income, gender, and household size are hypothesized to play a role not only in consumer preference for steak tenderness, but also in determining willingness-to-pay. Additionally, the amount of beef consumed, knowledge of the current USDA quality grading system, and preference for steak doneness may affect decisions.

Data and Results

Demographics and Consumption Habits

A total of 313 consumers participated in the study, 227 in the first treatment and 86 in the second.[8] Table 1 provides summary statistics for the two groups. A majority of participants were female and the average age in both groups was about 47 years. Participants on average had at least some college education and a household income between $40,000 and $50,000 in the first treatment and between $50,000 and $60,000 in the second treatment. Although the study includes only midwestern consumers, the participants

[7] "Guaranteed tender" and "Probably tough" steaks have slice shear force value of 15 kg or less and higher than 35 kg, respectively

[8] A fixed number of "in-store days" for conducting the experiments were allocated by grocery store administration The number of people that would enter the store each day was not known prior to the experiment, only the number of days in each store was known More store days were allocated to the first treatment of the experiment because the primary goal was to determine consumers' ability to identify tenderness and their willingness-to-pay for tenderness based on a blind taste test

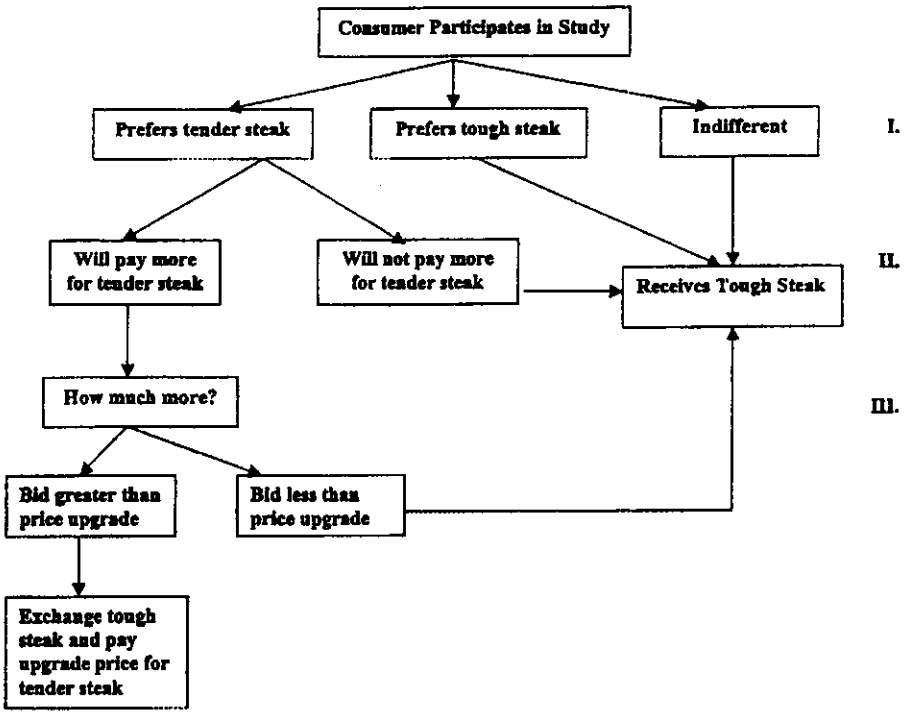

Figure 1. Consumer decision tree in steak tenderness valuation experiment

are representative of a wide range of demographics. For example, ages ranged from 19 to 82; education ranged from less than high school to Ph.D.; income ranged from less than $20,000/year to more than $120,000/year; and total beef consumption ranged from 0 times per week to over 12 times per week.[9]

Consumers on average ate ground beef about 2.2 times per week at home and slightly over 1 time per week away from home. Participants indicated that steak was consumed 1.1 times per week at home and 0.6 times per week away from home. Although the experiment was conducted near the meat counter, several participants that routinely consumed little or no beef took part in the study. Eighteen consumers indicated that they did not typically purchase beef at all and 60 consumers indicated that they did not typically purchase steak. Regarding the quality of beef typically purchased, 38% indicated they typically purchased USDA Choice, 19% indicated USDA Select, 19% indicated that they typically purchased store brand beef, and the remaining 22% did not know the grade of beef they usually purchased. Respondents, on average, preferred their steak cooked to a medium doneness.

Consumer Preferences for Sampled Steak

Table 2 shows the proportion of participants indicating preference for Red (guaranteed tender) or Blue (probably tough) steak in both treatments. In the first treatment, where consumers were not informed of the difference in the two steaks, the Red (guaranteed tender) steak was preferred overall by 69% of participants. Interestingly, of the four quality attributes, more people indicated they preferred the Red steak for tenderness (72%) than for any of the other attributes. This suggests consumers were able to determine independently that the primary distinguishing

[9] Participants in the experiments were more educated and had slightly higher incomes as compared to the general population. According to the U.S. Census Bureau (USCB), in 1998 the median household income was $38,885/year and 22.2% of the population had graduated from college. In our sample, the median yearly household income was between $40,000 and $49,999 and 41% of the respondents held at least a B.S. or B.A. degree. However, participant demographics were similar to the general population with respect to age and number of people living in the household. In 1998, the average age of the population (over 18) was about 46 and the average number of people living in the household was 2.63 (USCB). Experimental participants were on average 47.5 years of age and an average of 2.69 people lived in their households. A greater percentage of women participated in the in-store experiment (65%) than in the general population (51%), however, this difference is likely due to the fact that women traditionally do most of the shopping.

Table 1. Variable Definitions And Summary Statistics

Variable	Definition	Treatment Averages 1	Treatment Averages 2
Gender	1 if female; 0 if male	0.69 (0.47)	0.58 (0.49)
Age	Age of respondent in years	47.76 (16.06)	46.87 (12.91)
Education	Education level of respondent 1 = less than 12th grade; 2 = high school graduate; 3 = some technical, trade; business school; 4 = some college; 5 = B.S. B.A., complete; 6 = some graduate work; 7 = M.S., M.A. complete; 8 = Ph.D.. D.D.S., M.D., J.D., etc.	4.50 (1.88)	4.29 (1.60)
Income	Household income level 1 = less than $20,000; 2 = $20,000 to $29,000; 3 = $30,000 to $39,999; 4 = $40,000 to $49,999; 5 = $50,000 to $59,999; 6 = $60,000 to $69,999; 7 = $70,000 to $79,999; 8 = $80,000 to $89,999; 9 = $90,000 to $99,999; 10 = $100,000 to $109,999; 11 = $110,000 to $119,999; 12 = more than $120,000	4.39 (2.71)	5.78 (2.77)
Adults	Number of adults in household	1.96 (0.64)	2.02 (0.70)
Children	Number of children in household	0.71 (1.13)	0.74 (1.09)
Beef home	Number of times per week respondent consumes ground beef (hamburger) at home	2.28 (1.57)	2.21 (1.46)
Beef away	Number of times per week respondent consumes ground beef (hamburger) away from home	1.32 (1.58)	1.17 (1.21)
Steak home	Number of times per week respondent consumes steak at home	1.12 (0.85)	1.05 (0.82)
Steak away	Number of times per week respondent consumes steak away from home	0.54 (0.87)	0.63 (0.71)
Doneness	Preference of steak doneness; 1 = rare, 2 = medium rare, 3 = medium, 4 = medium well, 5 = well	3.22 (1.07)	3.05 (0.97)

Note: Number in parentheses are standard deviations. There were 227 participants in treatment 1 and 86 in treatment 2.

characteristic between the two steaks was tenderness.

In the second treatment, consumers were informed that one steak was guaranteed tender and the other was probably tough. When compared with results from the first treatment, consumers were more likely (84% compared to 69%) to indicate a preference for the Red or guaranteed tender steak. In fact, for all four quality attributes, the proportion indicating preference for the tender steak is higher when the differences in the steak samples were revealed. This suggests that labelling the steaks with relevant information influences revealed preferences.

Although a majority of respondents preferred the tender steak, many indicated they were not willing to pay more to exchange their tough steak for a tender one. In treatment 1, 69% preferred, but only 36% of the consumers were willing to pay extra to obtain a Red (guaranteed tender) steak. When tenderness was revealed in treatment 2, 84% preferred, but only 51% of the consumers were willing to pay for the guaranteed tender steak. The proportion of zero bids is surprisingly high, but there are several possible explanations. In particular, zero bids may be symptomatic of participants' prior experience with USDA grades in the sense that they lacked confidence that a second steak

Table 2. Respondent Steak Preferences

Attribute		Treatment 1 Red/Blue	Treatment 2 Tender/Tough
Taste	% preferring Red (Tender)	65.04	77.91
	% preferring Blue (Tough)	23.45[a]	10.47
Tenderness	% preferring Red (Tender)	72.12	89.53
	% preferring Blue (Tough)	18.58	4.65
Texture	% preferring Red (Tender)	65.48	69.77
	% preferring Blue (Tough)	22.12	13.95
Juiciness	% preferring Red (Tender)	59.73	75.58
	% preferring Blue (Tough)	27.88	10.47
Overall	% preferring Red (Tender)	69.16	83.72
	% preferring Blue (Tough)	22.02	9.30
	% willing to pay more for Red (Tender)	36.12	51.16
	Willingness to pay in $ per lbs[b]	1.23	1.84
Number of Participants		227	86

[a] Percentages do not sum to one in each category because some respondents indicated that they were "undifferent" between the two steaks with respect to the given attribute

[b] Average willingness to pay included only those individuals who offered a non-zero bid

from the same category would have similar qualities to the first (a consequence of the very problem that objective tenderness scoring is designed to address). That the proportion bidding zero was lower when steaks were identified as "probably tough" and "guaranteed tender" appears to lend support to this hypothesis.

Alternatively, zero bidding in this experiment may reflect the necessarily lower level of control in the grocery store setting compared to the lab. In particular, laboratory valuation studies typically require participants to be the consumers of the good. In this experiment the intended consumer may have been someone other than the participant—perhaps a spouse or other family member—in which case the preference revealed by the bid might not necessarily correspond to the stated preference based on the taste test. This will be an issue for future attempts to move experimental valuation to settings that are more familiar for the participant.[10]

Average willingness-to-pay (of those consumers who were willing to pay) was $1.23 per pound in treatment 1 and $1.84 per pound in treatment 2. The distribution of willingness-to-pay values for all consumers that preferred tender steak is presented in figure 2.

[10] Zero bids might also be due to cash constraints (i.e., the respondent may have only had a credit card or check and not cash), or time constraints (some participants probably did not want to take the time to complete the bidding portion).

Econometric Models

Figure 1 shows the three choices that consumers made in the experiment. First, respondents indicated which steak they preferred: Red (guaranteed tender) steak, Blue (probably tough) steak, or neither. A multinomial logit model was used to examine the impacts of various factors on the probability of steak choice (Greene). The model was estimated using 245 observations—less than the total number of participants (313) because some respondents failed to answer one or more survey questions. Marginal effects evaluated at the mean values are shown in table 3.

Age and education have a statistically significant positive influence on the probability that a consumer prefers tender steak. The estimates indicate that for every one-year increase in age, a participant is 0.4% more likely to prefer the tender steak. Thus, a 45 year old would be expected to be 8% more likely to prefer tender steak than would a 25 year old. For a one "unit" increase in the respondent's level of education (e.g., from "some college" to "B.S., B.A. complete"), a respondent is 4% more likely to prefer the tender steak. Perhaps the most notable result is the relatively large, positive, and statistically significant estimate for the "Treatment 2" dummy variable. Participants in the second treatment, where the tender and tough steaks were identified prior to tasting, were 18% more likely to prefer tender steak than participants in treat-

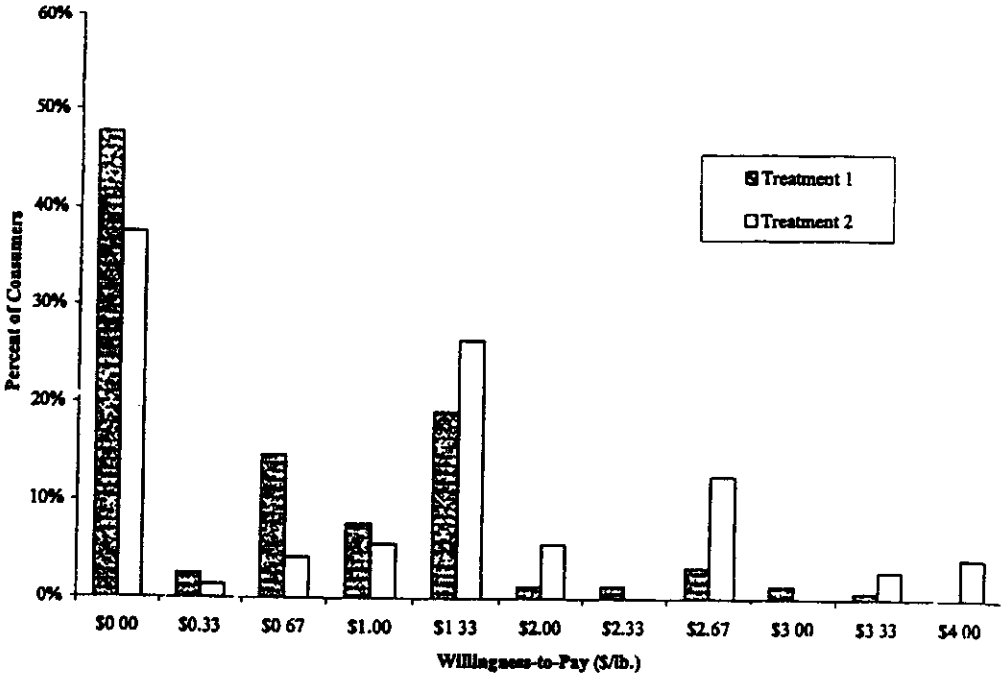

Figure 2. Distribution of willingness-to-pay premiums for tender steak

Table 3. Multinomial Logit–Marginal Probabilities of Factors Affecting Consumers Preference for Guaranteed Tender Steaks

	Steak preference category		
Variable	Tender	Tough	Neither
Age	0.004**	−0.003*	−0.001
	(0.002)	(0.002)	(0.001)
Education	0.040**	−0.027**	−0.013
	(0.014)	(0.012)	(0.008)
Gender	0.071	−0.027	−0.044
	(0.060)	(0.053)	(0.033)
Income	−0.005	0.007	−0.002
	(0.011)	(0.010)	(0.007)
Treatment 2	0.181**	−0.142**	−0.039
	(0.069)	(0.061)	(0.041)
Doneness	−0.044*	0.035	0.009
	(0.025)	(0.022)	(0.014)
Steak home	0.027	−0.047	0.020
	(0.034)	(0.032)	(0.016)
Steak away	−0.003	0.006	−0.003
	(0.034)	(0.029)	(0.019)
Knowledge of grade	−0.004	0.032	−0.028
	(0.066)	(0.059)	(0.035)

Note: Marginal effects evaluated at the means of the independent variables.
Number of observations = 245, log likelihood = −172.55, percentage of correct predictions = 74%.
* ** Denote significance at the 10% and 5% levels, respectively.
[a]Values in parentheses are standard errors.
[b]Knowledge of grade = 1 if respondent knew the quality grade of beef typically purchased, 0 otherwise.

ment one where tenderness levels were not identified. The only steak preference variable that influenced steak choice was cooking doneness – consumers who preferred their steaks cooked to a higher degree of doneness were less likely to prefer the tender steak. This is logical in that beef cooked to a higher degree of doneness is inherently tougher than beef cooked to a lower internal temperature (Milligan).

When respondents indicated a preference for the Red (guaranteed tender) steak, they were given the opportunity to bid to upgrade from their Blue (probably tough) steak to the Red (guaranteed tender) steak. With a significant proportion of the willingness-to-pay observations at the zero limit, an appropriate model for estimating willingness-to-pay is the tobit model. However, tobit estimation is restrictive in that it assumes that the probability of a limit outcome (a zero willingness-to-pay bid) and positive outcomes are affected identically by the same determinants. Haines, Guilkey, and Popkin showed that consumption of many food items often involves a two-step process. A common model to use in this instance is the double hurdle model developed by Cragg. Cragg's model is more general than the tobit in that it allows for different determinants of limit outcomes and positive outcomes.

Following Cragg, let P_i be the ith consumer's bid to upgrade to the Red (tender) steak. The first hurdle is the consumer's decision of whether to pay for the Red (tender) steak. If the consumer chooses not to bid, then $P_i = 0$. The probability of this outcome is

$$(1)\quad \text{Prob}(P_i = 0) = \Phi(-\boldsymbol{\beta}_1\mathbf{X}_i)$$

where Φ is the standard normal distribution function, $\mathbf{X}_i$ is a vector of consumer i's economic and demographic characteristics as previously indicated, and $\boldsymbol{\beta}_1$ is a vector of coefficients. The second hurdle determines the effect of the independent variables on P_i given that $P_i > 0$. Thus, the second hurdle involves the decision of how much to pay for the Red (tender) steak given that the consumer has decided to pay. The distribution of P_i conditional on being positive is truncated at zero and assumed normal with mean $\boldsymbol{\beta}_2'\mathbf{X}_i$ and variance σ_2. The second hurdle is formulated as

$$(2)\quad f(P_i|P_i > 0) = \{(1/\sigma)\phi(\lfloor P_i\boldsymbol{\beta}_2'\mathbf{X}_i\rfloor/\sigma)\} / \Phi(\boldsymbol{\beta}_2'\mathbf{X}_i/\sigma)$$

where ϕ is the standard normal density function and $\boldsymbol{\beta}_2$ is a vector of coefficients. Given equations (1) and (2), the likelihood for the random sample as shown by Haines, Guilkey, and Popkin is

$$(3)\quad L = \Pi\Phi(-\boldsymbol{\beta}_1'\mathbf{X}_i)^{(1-t_i)}\{\Phi(\boldsymbol{\beta}_1'\mathbf{X}_i)[(1/\sigma) \times\phi([P_i - \boldsymbol{\beta}_2'\mathbf{X}_i]/\sigma)]/\Phi(\boldsymbol{\beta}_2'\mathbf{X}_i/\sigma)\}^{t_i}$$

where t_i takes the value of 1 if $P_i > 0$ and 0 otherwise. This version of the double hurdle model assumes that the error terms in the two steps (the participation and consumption decisions) are independent and normally distributed. Results of the double-hurdle model are presented in table 4.[11]

The first column of results in table 4 shows the effects of consumer demographics on the probability that the consumer will pay for the Red (tender) steak. All variables are statistically insignificant at standard significance levels. The positive coefficient on "Knowledge of Grade" suggests that consumers who knew which quality grade of beef they typically bought may be more likely to pay for tenderness—but the effect is only marginally significant ($p = 0.21$). Although both income and treatment were expected to have positive effects, neither coefficient is statistically significant.

Estimates for the second hurdle indicate that females and younger consumers were willing to pay more for the upgrade to the tender steak. Surprisingly, income level did not significantly affect the amount respondents were willing to pay. The most important determinant of willingness-to-pay was the information treatment—consumers in the second treatment, where the tenderness levels were revealed, were willing to pay about $0.82 per pound more for tender steak and this effect was statistically significant ($p = 0.001$). The significance and magnitude of this estimate suggests that information and labelling have important economic impacts on consumer willingness-to-pay for a quality characteristic.

Implications and Conclusions

Current USDA quality grading standards do not provide consumers with a direct measure of tenderness. As a result, consumers

[11] A Lagrange Multiplier test, proposed by Lin and Schmidt, indicated that the double hurdle model was preferred to the tobit

Table 4. Double Hurdle Model—Determinants of Consumer Willingness-to-Pay for Guaranteed Tender Steak

Variable	First Hurdle Probability of Paying	Second Hurdle Amount Paid
Constant	0.275	1.478**
	(0.575)[a]	(0.613)
Age	−0.004	−0.016**
	(0.007)	(0.008)
Education	0.044	0.011
	(0.056)	(0.062)
Gender	−0.150	0.409*
	(0.210)	(0.237)
Income	−0.044	0.045
	(0.039)	(0.047)
Treatment 2	0.019	0.816**
	(0.212)	(0.258)
Household Size	−0.060	−0.083
	(0.089)	(0.091)
Steak home	0.044	0.028
	(0.121)	(0.135)
Steak away	0.101	−0.115
	(0.132)	(0.143)
Knowledge of grade	0.286	0.009
	(0.230)	(0.292)
Sigma[b]		0.911**
		(0.094)
Number of Observations	182	100

[a]Values in parentheses are standard errors

[b]Sigma is the error variance, log likelihood of first hurdle = −121.8, log likelihood of second hurdle = −110.2

*, ** Denote significance at the 10% and 5% levels, respectively

are uncertain when they purchase a steak (or other beef cut) whether it will be tender. In this experiment, participants tasted both tough and tender steaks and, if they preferred the tender steak, bid to upgrade from a tough to a tender steak. The willingness-to-pay values demonstrate some important points about consumer preferences for tenderness.

First, some consumers are willing to pay large premiums to obtain a "guaranteed tender" steak instead of a "probably tough" steak. Of the participants who were informed that the steak they were bidding for was "guaranteed tender," 20% were willing-to-pay a premium of $2.67 per pound or more. Second, providing consumers information via labelling has value. Participants in treatment 1 relied on a taste test alone to differentiate between the two steak samples. Participants in treatment 2 were told which steak was tender and which was tough. Given that information, more participants preferred the "guaranteed tender" steak (84% compared to 69%) and willingness-to-pay for the "guaranteed tender" steak increased by $0.82 per pound. Third, younger consumers and women were willing to pay significantly more to upgrade to a tender steak than were other consumers. Surprisingly, consumer income did not affect either the probability or the amount that participants would pay.

Overall, our results indicate that a quality-grading standard based upon (or supplemented with) tenderness measurements may be a viable alternative to the current system. Having established that the upgrade from a tough to a tender steak has value for consumers, future research is needed to explicitly compare the proposed tenderness based system with the current system—i.e., values for the upgrade from Select or Choice steaks to steaks of known tenderness—in order to better estimate the overall benefits.

Historically, the vast majority of beef has been marketed as a generic commodity with few branded products. Beef processors, who

for many years primarily focused on being low cost providers, are beginning to consider product differentiation and branding. Differentiation of beef via tenderness is beginning to occur among some industry participants, albeit at a slow rate. Information regarding the value of steak tenderness in addition to improvements in technology, that lower costs of tenderness identification, segregation, and alteration are the catalysts in this evolving market.

[Received October 1999; accepted August 2000.]

References

Becker, G.M., M.H. DeGroot, and J. Marschak. "Measuring Utility by a Single-Response Sequential Method." *Behav. Sci.* 9(July 1964):226–32.

Boleman, S.J. et al. "Consumer Evaluation of Beef of Known Categories of Tenderness." *J. Anim. Sci.* 75(June 1997):1521–4.

Boyce, R.R., T.C. Brown, G.H. McClelland, G.L. Peterson, and W.D. Schulze. "An Experimental Examination of Intrinsic Values as a Source of the WTA-WTP Disparity." *Amer. Econ. Rev.* 82(December 1992):1366–73.

Brester, G.W., and T.C. Schroeder. "The Impacts of Brand and Generic Advertising on Meat Demand." *Amer. J. Agr. Econ.* 77(November 1995):969–79.

Brester, G.W., and M.K. Wohlgenant. "Estimating Interrelated Demands for Meats Using New Measures for Ground and Table Cut Beef." *Amer. J. Agr. Econ.* 73(November 1991): 1182–94.

Buzby, J.C., J. A. Fox, R.C. Ready, and S.R. Crutchfield. "Measuring Consumer Benefits of Food Safety Risk Reductions." *J. Agr. Appl. Econ.* 30(July 1998):69–82.

Capps, O., D.S. Moen, and R.E. Branson. "Consumer Characteristics Associated with the Selection of Lean Meat Products." *Agribusiness* 4(November 1988): 549–57.

Chavas, J. "Structural Change in the Demand for Meat." *Amer. J. Agr. Econ.* 65(February 1983): 148–53.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica* 39(September 1971):829–44.

Dikeman, M.E. "Fat Reduction in Animals and the Effects on Palatability and Consumer Acceptance of Meat Products." *Proc. Recip. Meat Conf.* 40(1987):93.

Eales, J.S., and L.J. Unnevehr. "Demand for Beef and Chicken Products: Separability and Structural Change." *Amer. J. Agr. Econ.* 70(August 1988): 521–32.

——. "Simultaneity and Structural Change in U.S. Meat Demand." *Amer. J. Agr. Econ.* 75(May 1993): 259–68.

Flake, O.L., and P.M. Patterson. "Health, Food Safety and Meat Demand." Paper presented at the AAEA annual meeting, Nashville, TN, August 1999.

Fox, J.A., J.F. Shogren, D.J. Hayes, and J.B. Kliebenstein. "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *Amer. J. Agr. Econ.* 80(August 1998):455–65.

Greene, W.H. *Econometric Analysis.* Upper Saddle River, NJ: Prentice-Hall, 1997.

Haines, P.S., D.K. Guilkey, and B.M. Popkin. "Modeling Food Consumption Decisions as a Two-step Process." *Amer. J. Agr. Econ.* 70(August 1988):543–52.

Hayes, D.J., J.F. Shogren, S.U. Shin, and J.B. Kliebenstein. "Valuing Food Safety in Experimental Auction Markets." *Amer. J. Agr. Econ.* 77(February 1995):40–53.

Huffman, K.L., M.F. Miller, L.C. Hoover, C.K. Wu, H.C. Brittin, and C.B. Ramsey. "Effect of Beef Tenderness on Consumer Satisfaction with Steaks Consumed in the Home and Restaurant." *J. Anim. Sci.* 74(January 1996):91–7.

Kachelmeier, S.J., and M. Shehata. "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence for the People's Republic of China." *Amer. Econ. Rev.* 82(December 1992):1120–40.

Kinnucan, H.W., H. Xiao, C.-J. Hsia, and J.D. Jackson. "Effects of Health Information and Generic Advertising on U.S. Meat Demand." *Amer. J. Agr. Econ.* 79(February 1997):13–23.

Koohmaraie, M., C.N. Cutter, W.J. Dorsta, S.D. Shackelford, G.R. Siragusa, and T.L. Wheeler. "A Summary of Beef Quality and Meat Safety Research." Unpublished Report, USDA-ARS, U.S. Meat Animal Research Center, Clay Center, NE. 1996.

Lamb, R.L., and M. Beshear. "From the Plains to the Plate: Can the Beef Industry Regain Market Share?" Federal Reserve Bank of Kansas City, *Econ. Rev.*, Third Quarter 1998, Internet address: *http://www.kc.frb.org/publical/econrev/ermain.htm.*

LMIC. Livestock Marketing Information Center, Lakewood, Colorado. Internet address: *http://lmic1.co.nrcs.usda.gov*

Lin, T.F., and P. Schmidt. "A Test of the Tobit Specification Against an Alternative Suggested by Cragg." *Rev. of Econ. and Statist.* 66(February 1984):174–77.

Melton, B.E., W.E. Huffman, J.F. Shogren, and J.A. Fox. "Consumer Preferences for Fresh Food Items with Multiple Quality Attributes: Evidence from an Experimental Auction of Pork Chops." *Amer. J. Agr. Econ.* 78(November 1996):916–23.

Miller, M.F., K.L. Huffman, S.Y. Gilbert, L.L. Hammon, and C.B. Ramsey. "Retail Consumer Acceptance of Beef Tenderized with Calcium Chloride." *J. Anim. Sci.* 73(August 1995):2308–14.

Milligan, S.D., M.F. Miller, C.N. Oats, and C.B. Ramsey. "Calcium Chloride Injections and Degree of Doneness, Effects of Sensory Characteristics of Beef Inside Round Roasts." *J. Anim. Sci.* 75(March 1997):665–72.

Moon, W., and R.W. Ward. "Effects of Health Concerns and Consumer Characteristics on U.S. Meat Consumption." Paper presented at the AAEA annual meetings, Nashville, TN, August 1999.

Purcell, W.D. "Measures of Changes in Demand for Beef, Pork, and Chicken, 1975-1998." Research Institute on Livestock Pricing, Research Bulletin 3-98, Blacksburg, VA, October 1998.

Roosen, J., D.A. Hennessy, J.A. Fox, and A. Schreiber. "Consumers' Valuation of Insecticide Use Restrictions: An Application to Apples." *J. Agr. Resour. Econ.* 23(December 1998):367–84.

Savell, J.W., R.E. Branson, H.R. Cross, D.M. Stiffler, J.W. Wise, D.B. Griffin, and G.C. Smith. "National Consumer Retail Beef Study: Palatability Evaluations of Beef Loin Steaks that Differed in Marbling." *J. Food Sci.* 52(May/June 1987):512–17.

Shackelford, S.D., T.L. Wheeler, and M. Koohmaraie. "Tenderness-Based Classification of Beef." Unpublished manuscript, U.S. Meat Animal Research Center, USDA, Clay Center, NE. 1996.

Shogren, J.F., S. Cho, C. Koo, J. List, C. Park, P. Polo, and R. Wilhelmi. 2000. "Auction mechanisms and the measurement of WTP and WTA." *Resour. and Energy Econ.*, in press.

Shogren, J.F., S.Y. Shin, D.J. Hayes, J.B. Kliebenstein. "Resolving Differences in Willingness to Pay and Willingness to Accept." *Amer. Econ. Rev.* 84(March 1994):255–70.

Smith, G.C. et al. The National Quality Beef Audit. Colorado State University, Texas A&M University, College Station, and Oklahoma State University for the National Cattleman's Beef Association. 1995.

Unnevehr, L.J., and S. Bard. "Beef Quality: Will Consumers Pay for Less Fat?" *J. Agr. Resour. Econ.* 18(December 1993): 288–95.

U.S. Census Bureau. *http://www.census.gov.*

Vickrey, W. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *J. Finance* 16(1961):8–37.

Wheeler, T.L., L.V. Cundiff, and R.M. Koch. "Effect of Marbling Degree on Beef Palatability in *Bos taurus* and *Bos indicus* Cattle." *J. Anim. Sci.* 72(December 1994):3145–51