EXPERIMENTAL EVIDENCE OF RISK AVERSION IN CONSUMER MARKETS: THE CASE OF BEEF TENDERNESS

Frode Alfnes,[1] Kyrre Rickertsen[1] and Øydis Ueland[1,2]

[1]Norwegian University of Life Sciences

[2]Matforsk

Paper prepared for presentation at the

XIth International Congress of European Association of Agricultural Economists,

Copenhagen, Denmark, August 24-27, 2005

EXPERIMENTAL EVIDENCE OF RISK AVERSION IN CONSUMER MARKETS: THE CASE OF BEEF TENDERNESS

Frode Alfnes, Kyrre Rickertsen and Øydis Ueland

Date: July 14, 2005

Corresponding author: Frode Alfnes
Department of Economics and Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, 1432 Aas, Norway.
E-mail: frode.alfnes@umb.no

Frode Alfnes is a postdoctoral fellow in the Department of Economics and Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, N 1432, Aas, Norway. Kyrre Rickertsen is a professor in the Department of Economics and Resource Management, Norwegian University of Life Sciences and a senior research economist, Norwegian Agricultural Economics Research Institute, Oslo, Norway. Øydis Ueland is a researcher in Norwegian Food Research Institute, Aas, Norway. The corresponding author is Frode Alfnes.
The authors would like to thank employees at the Norwegian Food Research Institute for assistance with data collection and Eirik Mikkelsen for useful comments on an earlier version of this paper. Financial support was provided by The Research Council of Norway, Grant no. 147121/110.

EXPERIMENTAL EVIDENCE OF RISK AVERSION IN CONSUMER MARKETS: THE CASE OF BEEF TENDERNESS

Frode Alfnes, Kyrre Rickertsen and Øydis Ueland

Abstract
Consistency of quality is important for brand loyalty and market share in consumer markets. Among consumers of beef, tenderness is the primary quality attribute. We use an experimental auction market to investigate how inconsistency in tenderness affects consumers' willingness to pay (WTP) for beef. We find that both the level and the spread of tenderness affect consumers' WTP for beef. Categorization the beef into various classes of tenderness increased the total value of the beef by 8%, which suggests that improved tenderness labeling may be a profitable strategy.

JEL: C91, D12, D8, Q13.
Key words: beef tenderness, consumer demand, experimental auction, marketing, risk aversion.

Introduction
Risk preferences are important for consumers' choices in food markets, and several studies have already investigated consumers' willingness to pay (WTP) for food products associated with risk. These studies focus on either high stake but low probability risk, such as potential microbiological hazards resulting in illness or death (e.g., Hayes et al.,1995; Fox et al., 1998; Fox, Hayes, and Shogren, 1998), or food products that some consumers fear represent an unknown risk, such as genetically modified foods (e.g., Lusk, 2003; James and Burton, 2003) or hormone-treated beef (e.g., Alfnes, 2004; Alfnes and Rickertsen, 2003). The focus of this paper is risk aversion in consumer food markets with low stake but high probability risk. Consumers are exposed to this type of risk in the form of inconsistent quality attributes in food products. The probability of buying a tough steak is relatively high, for instance, but the negative consequences are generally less severe; a tough steak will not kill you, but it can ruin your dinner.

Holt and Laury (2002) investigated peoples' attitudes toward risk using lottery choices and found that risk aversion plays an important role in low, as well as high, stake choices. Their results indicate that risk aversion can plays an important role in the low stake choices food buyers make. However, a number of studies have found that risk preferences elicited in one context cannot be directly transferred another (e.g., Hudson, Coble, and Lusk, forthcoming; Hershey, Kunreuther, and Schoemaker, 1982). The importance of risk aversion in consumer markets then remains a largely unanswered empirical question.

Low stake and high probability risk is typical for unprocessed or semi-processed food products with inconsistent quality attributes. Producers of processed foods use a variety of quality control methods to ensure that products within a particular category have no, or barely any, detectable differences in important attributes. For unprocessed or semi-processed foods, quality control may include non-invasive techniques or the use of test samples, but these methods are often insufficient to accurately predict the quality of a batch because of the large biological diversity of natural produce. For example, two apparently similar cuts of beef can differ significantly in tenderness because of factors such as the animal's age, breed, gender, and stress level, and the tenderizing process employed. Thus, consumers are subject to much higher variability in the sensory experience when buying unprocessed or semi-processed foods, such as meat, fish, fruits, and vegetables, than when buying processed foods.

Tenderness is the primary determinant of satisfaction among beef consumers (Dransfield, Zamora, and Bayle, 1998; Sivertsen, Kubberød, and Hildrum, 2002), and inconsistent or inadequate tenderness are the top two beef-quality concerns for U.S. purveyors, restaurateurs, and retailers (National Beef Quality Audi, 2000). There are several studies of consumer preferences for beef tenderness. Acebron and Dopico (2000) investigated how consumers form in-store expectations about beef quality; Lusk et al. (2001) used experimental methods to examine consumer WTP for steak tenderness in a grocery store setting; and Lusk, Feldkamp, and Schroeder (2004) used several types of experimental auctions to elicit WTP for quality differentiated beef steaks. The results of these studies indicate that most consumers prefer tender beef and that many are willing to pay a premium for a guarantee of tenderness. However, none of these studies explored the effect of inconsistent tenderness on consumer WTP for beef. If, and as is hypothesized, consumers are risk averse, part of the premium paid for guaranteed tender beef may be due to the difference in tenderness variability between generic and guaranteed tender beef.

Market experiments with carefully controlled products, probabilities, and payoffs can be used to investigate consumers' WTP for various products and the importance of risk aversion in choices among products. There are several experimental methods for eliciting individual risk preferences; see Wakker and Deneffe for a discussion of risk preference elicitation methods. A common method is to elicit the WTP for risky prospects. The lower the WTP is relative to the expected value of the prospect, the higher the level of risk aversion (e.g., Harrison, 1986; Kachelmeier and Shehata, 1992; Pennings and Garcia, 2001).

We investigate how inconsistency in beef tenderness affects consumers' valuations by using an experimental auction to elicit consumers' valuation of beef in three tenderness categories as well as uncategorized generic beef. Given the known tenderness distribution of the generic beef, we calculate the ratio between the WTP for the generic beef and the expected value of the generic beef for each auction participant. The lower this risk ratio, the more risk averse the participant. The contributions of this paper are as follows. First, we show how lottery theory and experimental auctions can be used to study the effect of quality inconsistency on consumers' WTP for food products. Second, the effects of inconsistent beef tenderness on consumers' WTP for beef have not been studied in the past and we fill this empirical gap in the literature.

Experiment and Auction Design

In November 2003, we established a market experiment at the Norwegian Food Research Institute. We conducted four sessions with each session lasting approximately 90 minutes. Fifty-one participants were recruited through local organizations, including choirs and soccer teams, in southeastern Norway.[1] In each organization, the contact person was instructed to provide a sample of regular consumers of beef, between 25 and 60 years old, with an approximately equal division of sexes. Each participant was paid NOK 300[2] to participate. In addition, NOK 200 was paid to the recruiting organization for each participant who completed the experiment.

Table 1 presents the descriptive statistics for the participants. The participants' age ranged from 23 to 59 years, with an average of 38 years. Fifty-one percent were female. The average household income was NOK 531,000 and 55% had at least some university level education. The sample chosen is representative for the age group in the region.

We obtained 120 kilograms of beef sirloin from Norsk Kjøtt, Norway's largest meat processing company. The quality of the beef reflected the variation in quality found in local stores. The beef was produced on two consecutive days, vacuum-packed as whole loins, and stored for tenderizing for 14 days. The loins were then numbered from one to 48, cut into 1.5 cm slices, partitioned into portions weighing approximately 400 grams, and packed in consumer packages in a modified atmosphere similar to that found in Norwegian meatpacking.

Neither consumers nor experts can determine the tenderness of beef by visual inspection, and two cuts with the same marbling classification, for example U.S. select strip loin, can differ significantly in tenderness (Miller et al., 2001). We used the Warner-Bratzler Shear Force (WB) test to measure tenderness (Boleman et al., 1997). The WB test measures the amount of force required to penetrate a cut of meat: the lower the force required the more tender the meat. Consumer preferences regarding beef with various WB scores have been investigated in several papers (Boleman et al., 1997; Sivertsen, Kubberød, and Hildrum, 2002; and Huffman et al., 1996). Sivertsen, Kubberød, and Hildrum (2002), for example, concluded that WB test results are highly correlated with consumer ratings of tenderness with a sample correlation coefficient of –0.87.

The WB test was used to categorize the beef into three tenderness categories: very tender, tender, and less tender. This categorization was based on the measured distribution of the WB scores in the sample, i.e., the relative tenderness of the beef. Accordingly, the lowest 25% of WB scores in the sample were categorized as very tender, the next 50% as tender, and the remaining 25% as less tender.

Based on these tenderness categories, we created four qualities of beef: A, B, C, and V. Quality A was very tender, quality B tender, quality C less tender, and quality V uncategorized generic beef. In the experiments, we emphasized that quality V beef could be very tender, tender, or less tender and, furthermore, that the distribution of tenderness in quality V reflected the variation in the total beef sample, i.e., 25% very tender, 50% tender, and 25% less tender. Figure 1 was used to explain the tenderness variation of the four qualities. The participants were also shown samples to demonstrate that the various qualities did not differ in their visible attributes.

Several studies have used uniform fourth- or fifth-price auction to elicit WTP (e.g., Coursey, Hovis, and Schulze, 1988; Umberger and Feuz, 2004). Compared with the frequently used second-price auction the uniform fourth- (and fifth-) price auction has several benefits. First, a fourth-price auction will have a smaller difference between the average 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. Second, with multiple winners it is not as exclusive to win the auction and any auction winning utilities not associated with the product are reduced. Third, extreme outliers are less likely to affect the price information that the participants receive during the experiment. As the second-price sealed-bid auction, the uniform fourth-price sealed-bid auction is an incentive compatible mechanism for eliciting WTP (Vickrey, 1961). Bidding your WTP is then a weakly dominant strategy. Bidding more than your WTP gives a positive probability of winning when the price is higher than your valuation of the product. Bidding less than your WTP decreases the probability of winning when the price is less than the valuation (Shogren et al., 1994b).

We used a uniform fourth-price sealed-bid auction. In this type of auction, the participants submit sealed bids, the fourth-highest bid is the price, and the participants bidding higher than that price are the buyers. The participants were asked to bid simultaneously on the four qualities of beef. We started by explaining to the participants the auction procedure. Next, we conducted two hypothetical auction trials to familiarize the participants with the principles of the fourth-price sealed-bid auction. In these hypothetical trials, no real purchases were made. The participants were told explicitly that these trials were training trials that were to be followed by real trials. We then conducted four trials of non-hypothetical fourth-price sealed-bid auctions. We used multiple trials to make it possible to conduct within sample comparison of bids before and after introducing real economic incentives, before and after tasting the different qualities, and to allow the participants to refine their bids to more accurately reflect their own valuation. After each trial, the price for each quality and the identification number of the winning bidders were written on a whiteboard. After the four trials were completed, we randomly chose one binding trial and one binding quality. The participants with higher bids than the price for the binding quality in the binding trial were the buyers.

After two hypothetical and two real trials of bidding, the participants were allowed to taste samples of the different qualities. We informed the participants that we had randomly drawn one loin from each quality and that these loins were not necessarily representative of the respective qualities, but rather were examples on beef within the different qualities. We randomly drew one loin from each quality, sliced the loins into steaks, and grilled the steaks on both sides until the core temperature reached 60°C. Each steak was cut into small wedge shaped pieces, placed on white plastic dishes, and labeled with one of the letters A, B, C, or V depending on the quality. The same four loins were used in all four sessions. The loin drawn to represent V had a WB score categorizing it as very tender.[3]

Risk

The four qualities, A, B, C, and V, may be looked upon as lotteries, L_A, L_B, L_C, and L_V, where the outcome of each lottery is the tenderness of the beef. By segmenting the beef into tenderness categories, we have reduced but not eliminated the variability in tenderness for quality A, B, and C. The participants assign subjective probabilities for various tenderness outcomes for each of the four qualities. These subjective probabilities vary between the participants, and are based on the individual participant's experience and judgment. We have no information about the probabilities that the participants assign to the various tenderness outcomes within each quality, and neither do we have any information on their valuation on the various tenderness outcomes within each quality.

Each of the qualities A, B, and C may be viewed as a simple lottery, L_k, with k = A, B, C and where the outcome of each lottery is the tenderness of the beef, $t = 1,...,T$. Each simple lottery is a list $L_k = (p_1^k, ..., p_T^k)$ with $p_t^k \geq 0 \;\forall t$ and $\sum_t p_t^k = 1$, where p_t^k is the probability of tenderness *t* occurring in lottery *k*.

Quality V can be viewed as a simple lottery or as a compound lottery. A compound lottery, L^C, is the risky alternative that yields the simple lottery L_k with probability α_k. Quality V may be described as a compound lottery of the simple lotteries A, B, and C or $L_V^C = (L_A, L_B, L_C; \alpha_A, \alpha_B, \alpha_C)$ with $\alpha_k \geq 0 \;\forall k$ and $\sum_k \alpha_k = 1$. The compound lottery L_V^C can be reduced to the simple lottery $L_V = (p_1^V, ..., p_T^V)$, with the same ultimate distribution of tenderness probabilities. That is, the probability of outcome *t* in the reduced lottery is $p_t^V = \alpha_A p_t^A + \alpha_B p_t^B + \alpha_C p_t^C$. Whether the probabilities of various outcomes arise as a result of a simple lottery or of a more complex compound lottery has no significance, i.e., the simple lottery L_V is equally good as the compound lottery $L_V^C = (L_A, L_B, L_C; 0.25, 0.50, 0.25)$.

Assuming that a participant's preferences are represented by a von Neumann-Morgenstern utility function, the expected utility of the simple lottery L_V, $EU(L_V)$, is

(1) $$EU(L_V) = EU(L_V^C) = 0.25 \cdot U(L_A) + 0.50 \cdot U(L_B) + 0.25 \cdot U(L_C).$$

The first equality states that the simple lottery is equally good as the compound lottery and the second equality states that the expected utility is equal to the weighted average of the utilities of the simple lotteries L_A, L_B, and L_C with the probabilities of these lotteries as weights.

The certainty equivalent of a lottery, $CE(L)$, is the amount of money obtained with certainty that gives the same expected utility as the lottery. In an incentive compatible auction, a participant's bid for a lottery equals his or her *CE* of the lottery, or

(2) $$Bid(k) = CE(L_k).$$

The expected value of a lottery is the weighted average of the monetary values associated with the possible outcomes using the probability of each outcome as weights. In our case, the expected value of quality V, $EV(L_V)$, is the weighted average of the *CE*s for the qualities A, B, and C using the probabilities of each quality as weights

(3) $$EV(L_V) = EV(L_V^C) = 0.25 \cdot CE(L_A) + 0.50 \cdot CE(L_B) + 0.25 \cdot CE(L_C).$$

A risk-averse participant will bid less for a lottery than the expected value, a risk-neutral participant will bid the expected value, and a risk-seeking participant will bid more than the expected value. We calculate two measures of risk aversion. First, we calculate each participant's risk premium for the uncategorized generic beef by subtracting his or her bid for V from the expected value of V:

(4) $$Risk\ premium = EV(L_V) - Bid(\mathrm{V}).$$

Second, we calculate each participant's risk ratio by dividing his or her bid for V by the expected value of V:

(5) $$Risk\ ratio = Bid(\mathrm{V}) / EV(L_V).$$

A risk ratio of less than one implies risk aversion, a ratio equal to one implies risk neutrality and a ratio larger than one implies risk-seeking behavior (Di Mauro and Maffioletti, 2004).

Results

Tenderness

Table 2 presents the mean bids and their standard deviations. Columns H1 and H2 present the results of the two hypothetical auction trails. Columns R1 and R2 present the results of the non-hypothetical real trials conducted before tasting, while columns R3 and R4 present the results of the real trials conducted after tasting.

As expected, participants are willing to pay more for tender beef. In R1 and R2, the mean bid for A was 17% higher and the mean bid for C was 24% lower than the mean bid for B. In the two trials after the tasting, the mean bid for A was 21% higher and the mean bid for C was 19% lower than the mean bid for B. Furthermore, the mean bid for A, B, and C was 31%, 12%, and –15%, respectively, higher than the mean bid for V in R1 and R2, and 27%, 6%, and –14% higher in R3 and R4. All these bid differences are significantly different from zero with *p* values of less than 0.01 according to a Wilcoxon signed rank test[4]. The mean bids for quality B are significantly higher than the mean bids for quality V, even though they have the same median tenderness. Before tasting, the bid difference between B and V may partly be explained by the fact that the average valuation of A and C is 3.5% less than the valuation of B. However, after tasting, the average valuation of A and C is 1% higher than the valuation of B.

Tasting

We randomly drew one fillet from each quality for tasting. The fillet representing quality V would have been categorized as very tender (quality A). Consequently, the fillet representing quality B was less tender than the fillet representing quality V and tasting appears to have affected the bidding in R3 and R4. The mean bid for qualities A, C, and V increased by NOK 0.61 ($p = 0.29$), 2.67 ($p = 0.12$) and 0.88 ($p = 0.17$), respectively, from R2 to R3, whereas the mean bid for quality B decreased by NOK 2.84 ($p = 0.23$). None of these mean bids changed significantly between R2 and R3, although the mean bid for B was reduced compared with the bids for A ($p = 0.06$), C ($p = 0.01$), and V ($p = 0.02$). These changes indicate that the participants used the information obtained from tasting to update their beliefs about the four qualities, even though they were instructed that the loins were drawn randomly.

Rising bids

Figure 2 illustrates the rising mean bids throughout the trials. With the exception of a decline in the mean bid for quality B after tasting, the mean bids for all qualities are stable or rising. Rising bids in multi-trial auctions are one of the most persistent results in experimental valuation literature and are usually explained as learning effects. Framing and learning effects in multi-trial auctions have been studied in a number of papers (e.g., Knetsch, Tang, and Thaler, 2001; List and Shogren, 1999; Shogren et al. 1994a; Shogren, List, and Hayes, 2000). List and Shogren (1999) also found some evidence of bids being correlated with posted prices, but concluded that most of the increase in bids may be explained by participants learning their optimal strategy. The only study not reporting increasing bids is Knetsch, Tang, and Thaler (2001), who found that, in a ninth-price sealed-bid auction, bids were decreasing. Their results may be explained, at least partly, by the lack of competition in a market where almost all participants are buyers.

Bidding on all qualities simultaneously ensures that the rules and learning processes are identical for all qualities. Even though our bids are rising over the trials, the differences in bids between the different qualities seem to be robust estimators for the differences in WTP between these qualities.

Risk aversion

We compare the expected value, calculated as described by equation (3), to the bids for quality V in the six trials. In the first hypothetical trial, 12 participants bid more for V than the expected value of V. In the following five trials, 5, 6, 9, 12 and 18 participants bid the same or more for V than the expected value of V. Three participants (6%) bid the same or more for V than the expected value of V in all real trials, 25 participants (49%) bid the same or more in some trials and less in other trials, and 23 participants (45%) bid less for V than the expected value of V in all four real trials. In the four real trials, 13 participants (25%) bid on average more for V than the expected value of V. For the remaining 38 participants (75%), the average bid for quality V was lower than the expected value of V in the four real trials. These 38 participants showed evidence of risk aversion and were willing to pay less for the uncategorized beef than its expected value. Furthermore, the results from the four real trials imply that the categorization of the beef increased the total value of the beef by 8%.

As shown in table 3, the mean risk premiums were positive and the mean risk ratios were below one in all trials. The mean risk premium in R1 and R2 was NOK 7.00 and NOK 6.44 while it was reduced to NOK 4.96 and NOK 3.59 in R3 and R4. As discussed above, the tasting may explain this reduction. The results of the non-parametric Wilcoxon signed rank test show that the risk premium was statistically significant in all trials.

The cumulative distribution of the individual risk ratios in the four real trials is presented in figure 3. Two participants (4%) had a risk ratio below 0.7, five participants (10%) had a risk ratio below 0.8, 17 participants (33%) had a risk ratio below 0.9, 38 participants (75%) had a risk ratio below 1, and 49 participants (96%) had a risk ratio below 1.1.

Conclusions and Implications

In our experimental auction, 75% of the participants were risk averse in their choice of beef. The risk aversion evident suggests that consumers want to avoid inconsistencies in quality and, therefore, assign a negative value to uncertainty regarding tenderness. This result suggests that low stake but high probability risk is important in the beef market and is likely to be important in other markets for unprocessed or semi-processed foods such as meat, fish, fruits, and vegetables.

The participants were willing to pay 50% more for very tender beef and 25% more for tender beef compared with less tender beef. Further, the participants were not only interested in the average tenderness, but also the consistency of quality. On average, the bids for tenderness-categorized beef were 8% higher than the bids for non-categorized beef with identical tenderness distribution.

Currently, beef is sold as an experience good where the quality can be judged only after the purchase. Most beef is sold without any categorization of tenderness and beef processors have tended to focus on low-cost production. Labeling may transform beef from an experience good to a search good for which consumers have information regarding tenderness before purchase. In this study, the tenderness of the beef was accurately determined. Such precise determination may be beyond the scope of meat processors. However, the risk can be reduced by a tracking system that follows the production of beef. Factors that influence the end quality can then be monitored and a good estimate of tenderness provided as a basis for labeling. The profitability of such a marketing strategy depends on whether the tracking, labeling, and segmentation costs are lower than the predicted 8% increase in value.

Footnotes

[1] In global terms, Norway has a high organizational participation rate. In the Oslo area, for example, 49% of the population responds that they actively participate in at least one organization (Statistics Norway). Recruiting through organizations, therefore, gives a fairly representative sample of the population.

[2] November 14, 2003: NOK100 = US$14.31 = €12.20 (www.oanda.com)

[3] The WB scores for the sample fillets of A, B, C, and V were 31, 39, 58, and 32, respectively.

[4] We use this non-parametric test in all our bid comparisons.

References

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

Alfnes, F. "Stated Preferences for Country of Origin of Beef: Application of a Mixed Logit Model." *European Review of Agricultural Economics* 31(2004): 19–37.

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

Boleman S.J., S.L. Boleman, R.K. Miller, J.F. Taylor, H.R. Cross, T.L Wheeler, M. Koohmaraie, S.D. Shackelford, M.F. Miller, R.L. West, D.D. Johnson, and J.W. Savell. "Consumer Evaluation of Beef of Known Categories of Tenderness." *Journal of Animal Science* 75 (1997): 1521–24.

Coursey, D.L., J.L. Hovis, and W.D. Schulze. "The Disparity between Willingness to Accept and the Willingness to Pay Measures of Value." *The Quarterly Journal of Economics* 102(1988): 679–90.

Di Mauro, C., and A. Maffioletti. "Attitudes to Risk and Attitudes to Uncertainty: Experimental Evidence." *Applied Economics* 36(2004): 357–72.

Dransfield E., F. Zamora, and M.C. Bayle. "Consumer Selection of Steaks as Influenced by Information and Price Index." *Food Quality and Preference* 9 (1998): 321–26.

Fox, J.A., D.J. Hayes, and J.F. Shogren. "Consumer Preferences for Food Irradiation: How Favorable and Unfavorable Descriptions Affect Preferences for Irradiated Pork in Experimental Auctions." *Journal of Risk and Uncertainty* 24(2002): 75–95.

Fox, J.A., J.F. Shogren, D.J. Hayes, and J.B. Kliebenstein. "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *American Journal of Agricultural Economics* 80(1998): 455–65.

Harrison, G.W. "An Experimental Test for Risk Aversion." *Economics Letters* 21(1986): 7–11.

Hayes, D.J., J.F. Shogren, S.U. Shin, and J.B. Kliebenstein. "Valuing Food Safety in Experimental Auction Markets." *American Journal of Agricultural Economics* 77(1995): 40–53.

Hershey, J.C., H.C. Kunreuther, and P. Schoemaker. "Sources of Bias in Assessment Procedures for Utility Functions." *Management Science* 28(1982): 936–54.

Holt, C.A., and S.K. Laury. "Risk Aversion and Incentive Effects." *American Economic Review* 92(2002): 1644–55.

Hudson, M.D., K.H. Coble, and J.L Lusk. "Consistency of Risk Aversion Measures: Results from Hypothetical and Non-Hypothetical Experiments. " Forthcoming in *Agricultural Economics*.

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." *Journal of Animal Sciences* (1996): 91–7.

James, S., and M. Burton. "Consumer Preferences for GM Food and other Attributes of the Food System." *Australian Journal of Agricultural and Resource Economics* 47(2003): 501–18.

Johansson-Stenman, O., and H. Svedsäter. "Cheap Talk does not always come Cheap: Hypothetical Bias in Choice Experiments." Unpublished paper. Department of Economics, Göteborg University. 2004:1–29.

Kachelmeier S.J., and M. Shehata. "Examining Risk Preferences under High Monetary Incentives: Experimental Evidence for the People's Republic of China." *American Economic Review* 82(1992): 1120–40.

Knetsch, J.L, F.F. Tang, and R.H. Thaler. "The Endowment Effect and Repeated Market Trials: Is the Vickrey Auction Demand Revealing?" *Experimental Economics* 4(2001): 257–69.

List, J.A., and C.A. Gallet. "What Experimental Protocol Influence Disparities between Actual and Hypothetical Values?" *Environmental and Resource Economics* 20(2001): 241–54.

List, J.A., and J.F. Shogren. "Price Information and Bidding Behavior in Repeated Second-Price Auctions." *American Journal of Agricultural Economics* 81(1999): 942–49.

Lusk, J.L. "Effect of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice." *American Journal of Agricultural Economics* 85(2003): 840–56.

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

Lusk, J.L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. "In-store Valuation of Steak Tenderness." *American Journal of Agricultural Economics* 83(2001): 539–50.

Miller, M.F., M.A. Carr, C.B. Ramsey, K.L. Crockett, and L.C. Hoover. "Consumer Thresholds for Establishing the Value of Beef Tenderness." *Journal of Animal Science* 79 (2001): 3062–68.

Murphy, J.J., P.G. Allen, T.H. Stevens, and D. Weatherhead. "A Meta Analysis of Hypothetical Bias in Stated Preference Valuation." Forthcoming in *Environmental and Resource Economics,* 2004.

National Beef Quality Audit – 2000. "Improving the Quality, Consistency, Competitiveness and Market-Share of Fed-Beef. The Final Report of the Third Blueprint for Total Quality Management in the Fed-Beef (Slaughter Steer/Heifer) Industry". Retrieved October 4, 2004 from http://meat.tamu.edu/nbqa2000/nbqa2000.pdf.

Pennings, J.M.E., and P. Garcia. "Measuring Producers' Risk Preferences: A Global Risk Attitude Construct." *American Journal of Agricultural Economics* 83(2001): 993–1009.

Shogren, J.F., J.A. Fox, D.J. Hayes, and J.B. Kliebenstein. "Bid Sensitivity and the Structure of the Vickrey Auction." *American Journal of Agricultural Economics* 76(1994a): 1089–95.

Shogren, J.F., J.A. List, and D.J. Hayes. "Preference Learning in Consecutive Experimental Auctions." *American Journal of Agricultural Economics* 82(2000): 1017–21.

Shogren, J.F., S.Y. Shin, D.J. Hayes, and J.B. Kliebenstein. "Resolving Differences in Willingness to Pay and Willingness to Accept." *American Economic Review* 84(1994b): 255–70.

Sivertsen, H.K., E. Kubberød, and K.I. Hildrum. "Consumer Preferences of Beef Tenderness and Mechanical Measurements." *Journal of Sensory Studies* 17 (2002): 365–78.

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

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

Wakker, P., and D. Deneffe. "Eliciting von Neumann-Morgenstern Utilities when Probabilities are Distorted or Unknown." *Management Science* 42(1996): 1131–50.

Table 1. Descriptive Statistics for the Auction Sample

Variable	Definition	Mean[a]	St.dev.
Gender	Gender of participant		
	Male = 1; Female = 2	1.51	0.50
Age	Age of participant	38.29	8.50
Income	Total income of household[b]	5.31	1.89
	(in NOK 100,000)		
Education	Highest completed education	2.43	0.70
	Elementary school = 1		
	High school = 2		
	University/college = 3		

[a] Corresponding figures for the population between 20 and 60 years old in the Oslo area are 1.51, 39.80, 5.89, and 2.41, respectively, based on estimates from Statistics Norway.

[b] Income was categorized into six classes. The midpoints of each class were used in the estimations.

Table 2. Descriptive Statistics for the Bids

	Hypothetical Trials		Real Trials			
	H1	H2	R1	R2	R3	R4
Quality A						
Mean bid A	83.10	84.88	86.84	88.92	89.53	90.88
Standard deviation A	14.37	15.06	13.71	12.10	13.42	12.33
Price premium A[a]	39%	38%	34%	33%	33%	28%
Quality B						
Mean bid B	69.16	72.53	73.88	76.92	74.08	75.57
Standard deviation B	12.67	14.01	14.24	12.25	15.06	15.03
Price premium B	16%	17%	13%	14%	9%	6%
Quality C						
Mean bid C	54.22	56.22	57.41	57.37	60.04	61.14
Standard deviation C	13.38	14.10	14.08	16.21	16.20	12.34
Price premium C	–10%	–10%	–12%	14%	–12%	–14%
Quality V						
Mean bid V	61.16	63.49	66.00	68.59	69.47	72.20
Standard deviation C	13.98	12.98	12.81	13.54	14.31	14.47

[a] Price premium for quality A relative to quality V, measured in percent. Price premium A = 100*(Bid A – Bid V)/Bid V.

Table 3. Preferences toward Risk

	Hypothetical Trials		Real Trials				
	H1	H2	R1	R2	R3	R4	Mean
Mean risk premium	7.75	8.04	7.00	6.44	4.96	3.59	5.50
Standard deviation	10.64	7.45	7.50	7.84	9.75	10.19	6.96
Wilcoxon test *p* value[a]	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Mean risk ratio	0.89	0.89	0.91	0.91	0.93	0.96	0.93
Standard deviation	0.15	0.12	0.10	0.11	0.14	0.15	0.10
# Risk averse	38	43	43	39	35	32	38
# Risk neutral	1	3	2	3	4	1	0
# Risk seeking	12	5	6	9	12	18	13

[a] Wilcoxon signed rank test. H_0: Mean risk premium equal zero.

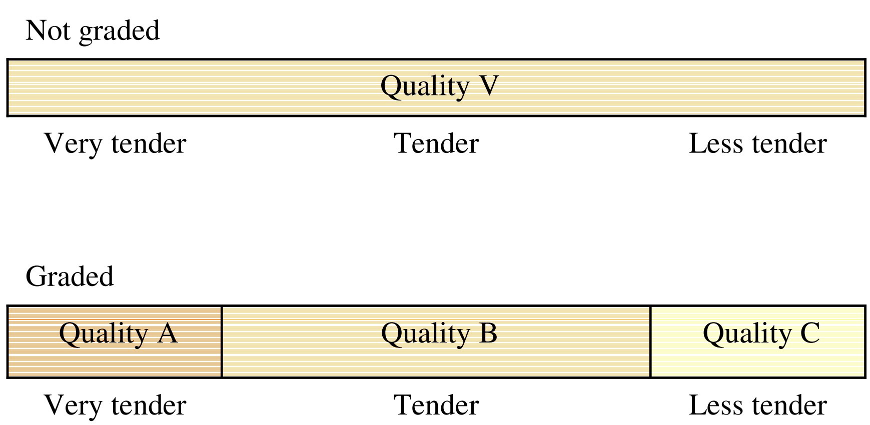


Figure 1. Tenderness Grading

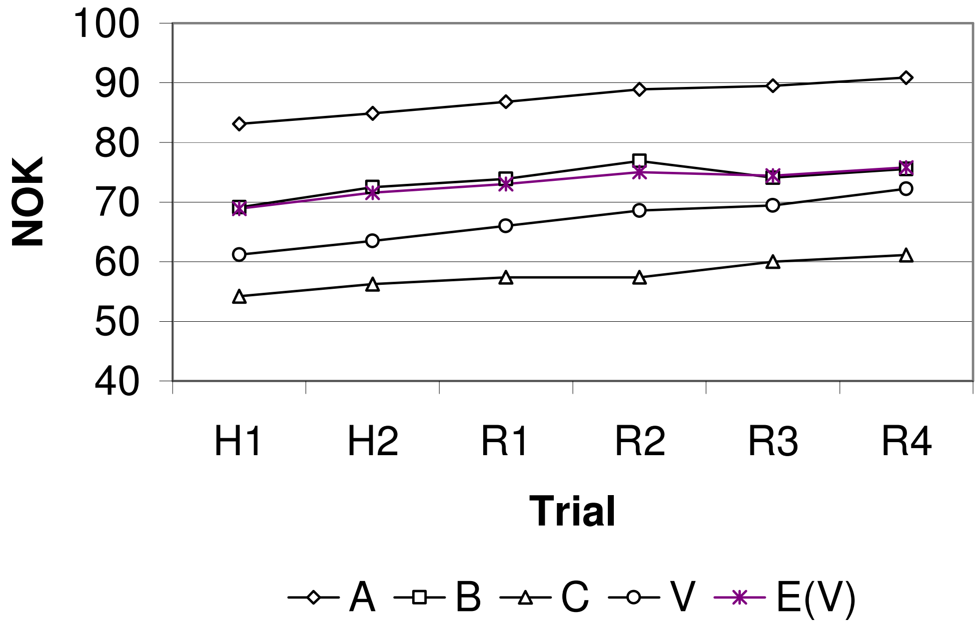
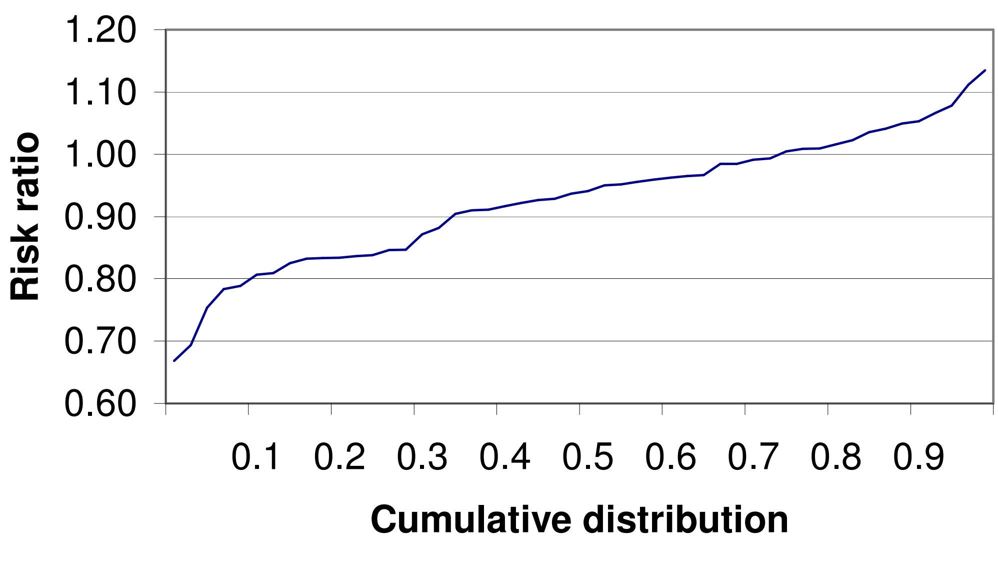


Figure 2. Willingness to Pay for the Different Qualities



[a] The risk ratio is defined as the WTP for generic beef divided by the weighted average of the WTP for the three qualities of tenderness-categorized beef, with the shares used in the categorization as weights.

Figure 3. Cumulative Distribution of the Risk Ratio[a]