# Preference structures and behavioral consistency in negotiations [1]

Rudolf Vetschera
University of Vienna
Austria
Rudolf.Vetschera@univie.ac.at

## Abstract

In this paper, we report results of the analysis of about 4,700 multi-attribute utility functions elicited during experiments with the Internet-based Negotiation Support System Inspire. The empirical results indicate that common assumptions of decision analysis, like monotonicity of single-attribute utilities or decreasing marginal utilities, are violated in a significant number of cases. Nevertheless, many structural properties of utility functions are clearly reflected in the outcomes of negotiations. On the other hand, behavior during the negotiation process contradicts the preferences implied in the utility functions in about 25% of all cases.

[1]Paper presented at the MCDM world conference, Whistler Mountain, Canada, August 6-11, 2004. I would like to thank Vivi Nastase for extracting the Inspire negotiation records. This research is part of the eNegotiation project supported by the Social Sciences and Human Ressources Council, Canada.

# 1 Introduction

Empirical research in the context of multi-criteria decision making has mainly focussed on two types of studies. On one hand, empirical studies were conducted to compare different methods for solving multi-criteria decision problems. The majority of these studies was performed in the area of multi-attribute utility theory and related approaches, where different methods for eliciting attribute weights or single attribute utility functions were compared via experiments (Schoemaker & Waid, 1982; Borcherding, Eppel, & Winterfeldt, 1991; Kimbrough & Weber, 1994; Delquie, 1997; Wang & Yang, 1998; Fischer, Carmon, Ariely, & Zauberman, 1999; Beroggi, 2000). Other experiments compared utility based methods to other approaches (Corner & Buchanan, 1995; Buchanan & Corner, 1997). On the other hand, empirical studies were carried out to explore the impact of various bias phenomena that occur in multi-criteria decision problems, like range effects (Nitzsch & Weber, 1993), the splitting bias in assessing weights (M. Weber, Eisenführ, & Winterfeldt, 1988), or reference and anchoring effects (Delquie, 1993; Buchanan & Corner, 1997).

Most of these studies were carried out in an experimental setting using a rather limited sample of subjects. Only a few studies used a larger sample size of a few hundred subjects (Borcherding et al., 1991) or were carried out as field research involving several hundred respondents (Hadley, Schoner, & Wedley, 1997).

The present paper poses two somewhat different and in a way more fundamental research questions. It is based on a quite large database of preference assessments, which were collected in negotiation experiments performed with the experimental negotiation support system Inspire on the Internet. This database allows us to study whether utility functions, which are elicited from users who are not specifically trained in decision analysis, actually correspond to standard assumptions usually made in decision analysis, like monotonicity of single attribute utility functions, or decreasing marginal benefits.

This question is important for several application. Firstly, several multi-criteria decision making methods implicitly rely on these assumptions. For example, the notion of dominance in criteria space as it is usually defined only makes sense when preferences are monotonic in each attribute. Secondly, in many studies multi-criteria decision making methods are compared using simulation experiments (Triantaphyllou & Mann, 1989; Fry, Rinks, & Ringuest, 1996; Stewart, 1996; Zanakis, Solomon, Wishart, & Dublish, 1998) in which decision makers are simulated by a utility function exhibiting "standard" properties. If it turns out that empirically, many utility functions do not exhibit these properties, the results of these simulation experiments for real decision makers must be put into question.

Our second research question relates these utility functions to actual behavior. In the Inspire system, utility functions are used to provide negotiators with an evaluation of offers made by their opponents and by themselves. But the system does not enforce a particular behavior, so subjects are free to behave consistently with their utility functions or not. It is therefore possible to study whether the utility ratings (or particular structural features of the utility functions) are reflected in the subjects' actual behavior.

The present study also differs from other experiments in the size and composition of the data used. In the time frame of 1996-2004, on which this paper is based, several thousand negotiation experiments were carried out and more than 6,000 utility elicitations were performed. This leads to a sample which is considerably bigger than in previous empirical studies in the field of multi-criteria decision making. However, since the system Inspire is openly available on the Internet, the user population underlying this study is less well defined than in controlled laboratory experiments. The present paper is thus more exploratory in nature and should be considered as field research rather than laboratory experiments.

The remainder of this paper is structured as follows: in section two, we introduce the Inspire Negotiation Support System, the case used, and provide an overview of the empirical data which forms the basis of our analysis. In section three, we focus on different structural properties of the utility functions elicited in Inspire and analyze how closely they relate to standard properties like monotonitcy and whether these structural properties are reflected in the negotiator's behavior or the outcome of negotiations. Section four compares the actual behavior of subjects to the behavior prescribed by the utility models as a whole. Section five concludes the paper by summarizing its results, discussing their relevance for preference modeling, and providing an outlook on future research.

# 2 The Inspire Database

## 2.1 Inspire negotiations

The Inspire Negotiation Support System (Kersten & Noronha, 1999) has been available on the Internet since 1996. It is an experimental system used for teaching and research on computer supported negotiations. A large part of its users are students, who carry out negotiations as part of courses on negotiation, decision analysis, information systems or similar subjects. But the system is freely available on the Internet and participation in negotiation experiments is open to the general public, too.

Inspire is a web-based system which provides several tools to support the entire negotiation process. It contains a communication platform through which negotiators can exchange structured offers addressing all issues (attributes) of the problem being negotiated, as well as unstructured text messages. The analytical support tools of Inspire provide methods for eliciting the utility functions of negotiators (which are private information of the parties), tools to evaluate offers from both sides, graphical displays of the negotiation dynamics and tools to identify and improve upon dominated (not Pareto-efficient) compromises.

Negotiation experiments in Inspire follow a structured pattern. Users enter a negotiation experiment (or a set of experiments which is set up, for example, for a course at a particular university) by defining a negotiation name for themselves to preserve anonymity in the experiments. Users are then assigned a role in the experiment and matched with a partner playing the complementary role (e.g. buyer and seller). The negotiation case is presented to the user and their utility function for the attributes relevant to the case is elicited using a conjoint measurement method. Typically, the case description specifies the attributes which are relevant for the case and the direction of improvement (for example, that a seller in a buyer/seller negotiation should prefer a higher price over a lower price), while specific utility values and trade-off weights between attributes are elicited from the users. Users can change their utility functions by performing another elicitation at any time during the negotiation.

After the utility function elicitation is completed, the users fill in a pre-negotiation questionnaire, in which demographic data about the users and a subjective rating of the utility function elicitation and their understanding of the case are recorded. Then the two partners start to negotiate by exchanging offers and/or text messages. The system automatically calculates and displays the utility values for all offers, and users can request additional graphical displays. Usually, a time limit of three weeks is set for the negotiation. When the two parties have found a compromise, the system determines whether the compromise is Pareto-optimal. If it is not, the system proposes alternatives which dominate the compromise and the negotiation continues. When the negotiation is completed, a post-negotiation questionnaire is administered to the users, in which they are asked about their perceptions and assessment of the negotiations.

Inspire can be used for different negotiation cases. However, most of the negotiation experiments carried out so far used a single case, the "Cypress/Itex"-case (written by Dr. David Cray from Carleton university). Only experiments based on this case are used for the analyses performed in this paper. The Cypress/Itex-case is a buyer/seller negotiation about the purchase of bicycle parts. The two parties negotiate about four attributes of a purchasing contract: the price of the parts, delivery time, terms of payment and the return policy for defective parts. For each attribute, a set of possible values is pre-defined in the case. There are five possible values for price, four for delivery time and three each for payment and return policy. Thus, there is a total of 180 alternative contracts from which the parties can choose. Within all attributes, the preferences of the two parties are strictly conflicting. However, since the attribute weights and the utility functions within the attributes are specified by the negotiators themselves, it is possible that some alternatives are dominated in utility space.

## 2.2 The Inspire database

The Inspire systems records all data generated during the negotiation experiments. The database utilized for this paper contains data from 2,814 experiments based on the Cypress/Itex case. For each experiment, the pre- and post negotiation questionnaire, all utility functions elicited from users and all offers and messages exchanged during the negotiations are stored.

Demographic data from the pre- and post negotiation questionnaires is available from 5,625 users from 74 countries. To allow for an analysis of cultural effects, the analysis was restricted to data from countries with more than 50 users and to users who were born and live in the same country. The resulting geographical distribution of users is shown in table 1, table 2 provides an overview of the original and the reduced data set.

| Country | | $N$ |
|---|---|---|
| Austria | AT | 171 |
| Canada | CA | 493 |
| Switzerland | CH | 51 |
| Germany | DE | 613 |
| Ecuador | EC | 195 |
| Finland | FI | 168 |
| Hong Kong | HK | 85 |
| India | IN | 342 |
| Norway | NO | 68 |
| Poland | PL | 53 |
| Russia | RU | 323 |
| Taiwan | TW | 145 |
| USA | US | 610 |
| Total | | 3,317 |

Table 1: Number of users from countries included in the analysis

The pre-negotiation questionnaire provides several demographic variables about users, which are summarized in table 3. Figure 1 shows the age distribution of users, the average age was 26.18 years, the median 23.63 years, so the distribution is somewhat skewed.

As can be seen from table 2, there are more utility functions than users, so there is a number of users who performed several utility elicitations. Table 4 shows the distribution of the number of utility estimations across users. Over three quarters of all users performed

| | all data | used |
|---|---|---|
| Total experiments | 2,814 | 2,292 |
| Number of users | 5,625 | 3,317 |
| Negotiations started | 2,645 | 2,261 |
| Compromise reached | 1,586 | 1,114 |
| Users for which utility functions are available | 4,835 | 3,307 |
| Number of utility functions elicited | 6,892 | 4,747 |

Table 2: Summary of data used

only one utility elicitation, the remaining ones changed their utility function only a few times.

Figure 2 shows the time intervals between subsequent utility estimations. A considerable number of changes to the utility functions took place rather quickly, perhaps to correct a mistake that has been made. On the other hand, there is also a significant number of users who changed their preferences well into the negotiation process, which could be an indication of an instability of preferences.

## 3 Preference Patterns

In this section, we will analyze different characteristics of the utility functions elicited from Inspire users and identify factors which could cause the observed deviations of these functions from standard assumptions of decision analysis.

As was already mentioned, discrete values are used for all attributes in negotiation problems in Inspire. The conjoint measurement method used to elicit utility functions assigns a partial utility value to each of those discrete values. To relate the utility functions elicited in Inspire more closely to standard assumptions of decision analysis, we express them in terms of single attribute utility functions for each attribute and attribute weights as:

$$u(\mathbf{x}) = \sum_k w_k u_k(x_k) \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_K)$ is the vector of attribute values describing alternative $\mathbf{x}$, $w_k$ is the attribute weight and $u_k(.)$ is the single attribute utility function for attribute $k$.

Let $v_{kj}$ be the utility assigned by the conjoint measurement method to the $j$-th discrete value of attribute $k$. The $v_{kj}$ are automatically scaled by the elicitation method so that $\min_j \ v_{kj} = 0$, i.e. the worst outcome in each attribute is assigned a utility of zero.

| Variable | Scale | Explanation | |
|---|---|---|---|
| Gender | Nominal | Gender of the user | Female: 42.18%<br>Male: 57.82% |
| Country | Nominal | Country in which user was born and lives | |
| Role | Nominal | Role (Buyer/Seller) in the negotiation experiment | Buyer: 46.06%<br>Seller: 53.94% |
| NSS Before | True/False | Indicates whether the user has used an NSS before | yes: 11.23%<br>no: 88.77% |
| Nego. Experience | Likert | Self-rated experience in negotiations,<br>1 = Very experienced,<br>5 = No experience | M=3.61<br>SD=1.04 |
| Weight Issues | Likert | How easy was it to weight issues,<br>1 = Extremely easy,<br>5 = Extremely difficult | M=2.83<br>SD=0.91 |
| Weight Options | Likert | How easy was it to weight options,<br>1 = Extremely easy,<br>5 = Extremely difficult | M=2.96<br>SD=0.94 |
| Understand Case | Likert | How understandable was case to user,<br>1 = Extremely easy,<br>5 = Extremely difficult | M=2.18<br>SD=0.83 |
| Age | Numeric | User's age in years | M=26.18<br>SD=6.82 |

Table 3: Overview of socio-demographic variables used

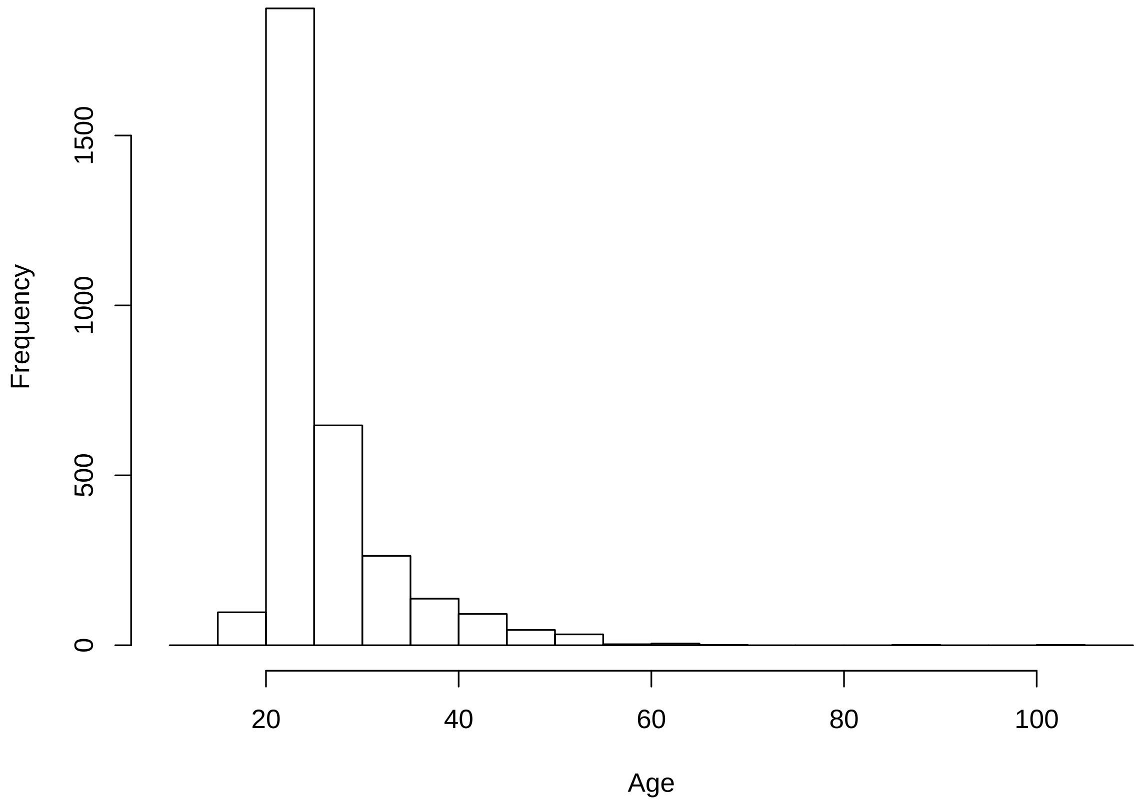


Figure 1: Distribution of the users' age

Furthermore, the optimal alternative for a decision maker is assigned a utility of 1. The weights $w_k$ can then simply be inferred from the values $v_{kj}$ as

$$w_k = \max_j \; v_{kj} \tag{2}$$

By dividing all $v_{kj}$ by the corresponding $w_k$, we also obtain values of the single attribute utility functions $u_k(.)$ for the different attribute values scaled in the interval [0,1].

In the following sections of the paper, we will analyze the following structural properties of utility functions :

- Monotonicity of utility functions for single attributes
- Attribute weights
- The curvature (convexity or concavity) of single attribute utility functions

| Elicitations | Percent |
|---|---|
| 1 | 76.75 |
| 2 | 14.15 |
| 3 | 4.63 |
| 4 | 2.06 |
| 5 and more | 2.42 |

Table 4: Distribution of number of utility elicitations

### 3.1 Monotonicity

The case description states that sellers should have a preference for higher prices, longer delivery times, shorter terms of payment and a higher possible rate of defective parts, and buyers vice versa. Thus, sellers should exhibit strictly monotonic increasing utility functions for the attributes price, delivery time and returns, and strictly monotonic decreasing utilities for terms of payment.

To analyze this property, all single attribute utility functions were classified into five categories according to their direction of improvement and whether they were strict monotonic, monotonic or not monotonic at all. Table 5 shows the distribution of these five categories for buyers, sellers and all subjects. When aggregating data from buyers and sellers, the opposite directions were taken into account, so the first column in category "all" refers to utility functions which were strict monotonic in the appropriate direction for the respective role.

| Attribute | | Strictly decreasing /correct | Decreasing /correct | Not Monotonic | Increasing /incorrect | Strictly increasing /incorrect |
|---|---|---|---|---|---|---|
| Price | Buyer | 68.39 | 18.46 | 11.59 | 0.15 | 1.41 |
| | Seller | 1.95 | 1.34 | 21.69 | 17.13 | 57.88 |
| | All | 63.13 | 17.80 | 16.64 | 0.74 | 1.68 |
| Delivery | Buyer | 72.09 | 16.80 | 9.30 | 0.34 | 1.46 |
| | Seller | 5.29 | 5.05 | 23.40 | 15.06 | 51.20 |
| | All | 61.65 | 15.93 | 16.35 | 2.69 | 3.38 |
| Payment | Buyer | 8.62 | 1.70 | 10.23 | 6.72 | 72.72 |
| | Seller | 79.32 | 7.04 | 9.04 | 0.61 | 3.99 |
| | All | 76.02 | 6.88 | 9.63 | 1.16 | 6.31 |
| Returns | Buyer | 81.44 | 10.18 | 5.80 | 0.93 | 1.66 |
| | Seller | 5.01 | 1.42 | 9.52 | 7.16 | 76.88 |
| | All | 79.16 | 8.67 | 7.66 | 1.17 | 3.33 |

Table 5: Distribution of different types of monotonicity (in %)

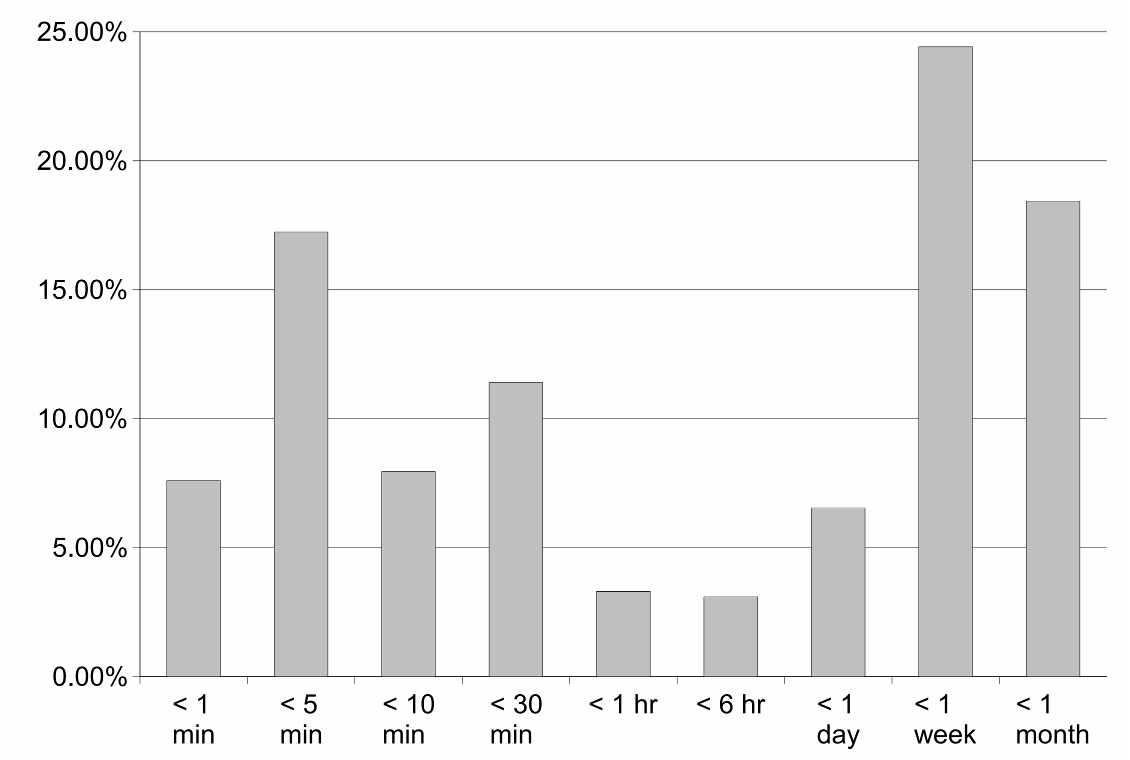


Figure 2: Time interval between utility estimations

While in all cases an overwhelming majority of users exhibited strict monotonic or at least monotonic single attribute utility functions in the right direction for their respective role, there was nevertheless also a considerable number of non-monotonic functions and even monotonic functions in the wrong direction.

There are three possible explanations for these deviations from the correct encoding of preferences:

- These are simply encoding errors, which were subsequently detected and corrected by users;
- the users understood the case correctly, but were not able to represent their preferences in terms of a utility function; or
- the users misunderstood the case, and the utility function correctly represents their wrong preferences.

The first explanation seems quite plausible in view of the fact that subsequent utility elicitations were sometimes performed in rather short intervals, as indicated in figure 2. To check whether these quick re-estimations of utility functions really served to correct errors in monotonicity, we compare the fraction of single attribute utility functions with wrong monotonicity among those utility functions which were changed within 15 minutes to those cases in which no such change took place.

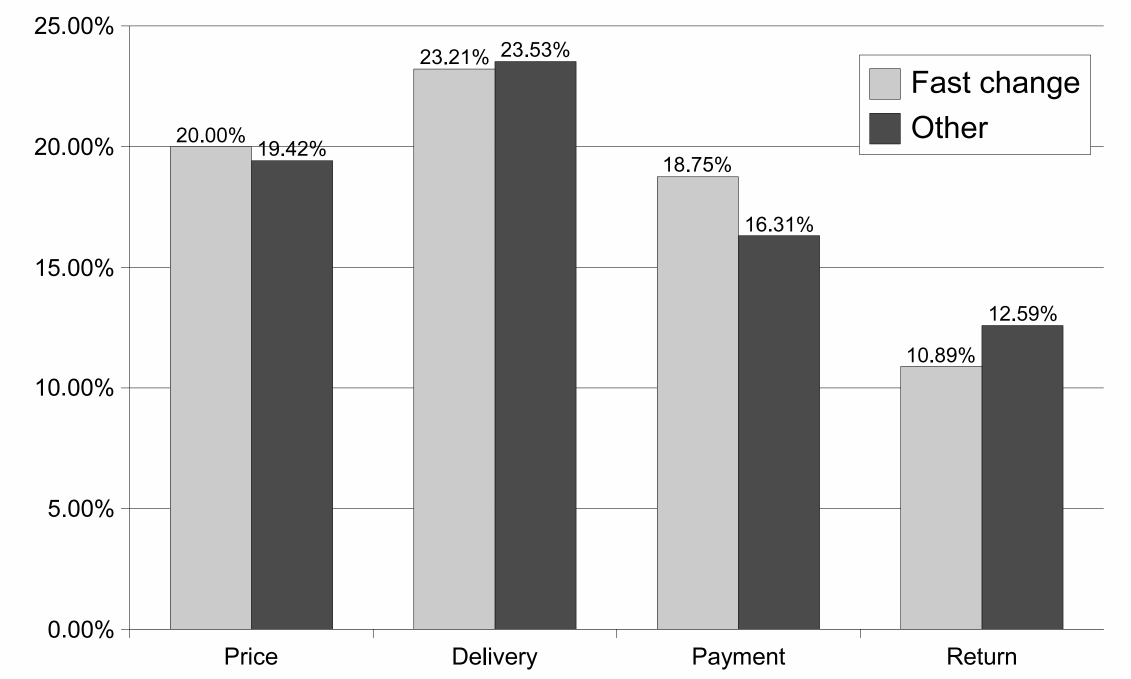


Figure 3: Monotonicity errors in rapidly changed utility functions

As figure 3 shows, fast changes in utility functions do not necessarily lead to an improvement in terms of monotonicity errors. In some cases, for example for the attribute returns, utility functions which are changed exhibit even fewer errors than functions which were retained for a longer time. A $\chi^2$ test indicated that for none of the attributes the differences are statistically significant. Thus we can reject the hypothesis that errors in monotonicity were quickly detected and corrected by users.

This leaves the other two explanations, either that correct preferences were wrongly encoded in the utility functions, or that the users' preferences as such were wrong due to a misunderstanding of the case. These two effects can be distinguished by considering the outcomes of the negotiations. If for example in a negotiation, the seller mistakenly tries to minimize the price, and the buyer (who has correct preferences) does the same, then the outcome should be the minimum price.

Figure 4 clearly shows that the utility functions in most cases reflected the actual preferences the users were following in the negotiations. Whenever both parties tried to minimize or maximize the same attribute according to their respective utility functions, the result in most cases indeed was the appropriate extreme value. Thus we can conclude that even users who thoroughly misunderstood the case were frequently able to represent their (incorrect) preferences in the form of a single attribute utility function.

Table 6 shows the results of a logistic regression analysis to identify possible factors which influence the probability of making an error in monotonicity. The table provides the parameter estimates and corresponding values of the $\chi^2$ test. Table 7 contains the odds ratios which indicate the change in the probability of not making an monotonicity error caused by the various factors. Effects which are significant at the 1 % level are marked with two asterisks (**), effects significant at the 5 % level with one asterisk.

Two variables have a consistent influence on monotonicity errors in all attributes: the role of a user in the experiment (buyer or seller), and their understanding of the case, which was encoded on a 5 point Likert scale from 1=extremely easy to understand to 5= extremely difficult to understand. Compared to sellers, buyers are about half as likely to make a monotonicity error in the attributes price and returns, and even less likely (about 22 %) to make an error in the attribute delivery time, but they make significantly more errors in the attribute payment conditions. This phenomenon corresponds with the direction in which the attributes are to be optimized, it seems that attributes which should be minimized lead to less errors in monotonicity than attributes which are to be maximized.

The consistent influence of the users' understanding of the case confirms our analysis above: the less well a user understood the case, the higher is the user's probability of making an monotonicity error in specifying the utility function.

A weakly significant effect is caused by the user's age, where older users in general have a higher probability of making a monotonicity error. This effect could also be related to the occupation of users, since students (who are in general younger) might be more familiar with the concept of utility functions than other users. There are also significant effects of the users' country of residence, but they do not follow clear-cut cultural patterns.

## 3.2 Weights

Figure 5 gives an overview of the distributions of the weights assigned by users to the four attributes. As can be seen from this figure, most users assigned a weight which is about twice as high to prices (Average = 0.3893) than to the other attributes (Average = 0.2012, 0.2069 and 0.2026, respectively).

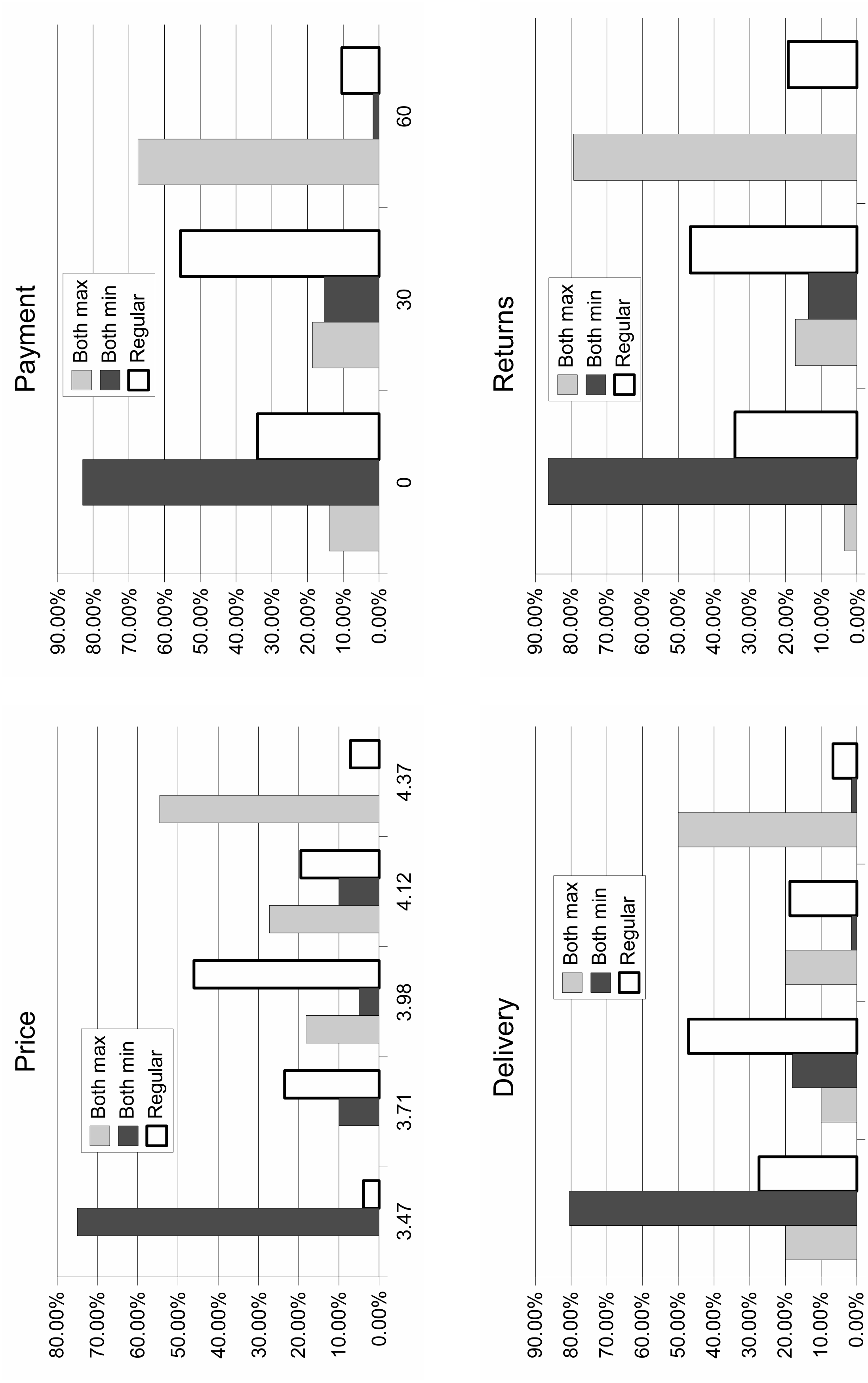


Figure 4: Effect of monotonicity errors on compromise

| Variable | | Price | Delivery | Payment | Returns |
|---|---|---|---|---|---|
| Gender | F | 0.0475 | -0.0576 | -0.0353 | -0.0025 |
| | | (0.8426) | (1.3603) | (0.4108) | (0.0017) |
| Country | AT | -0.2040 | *-0.6029 | 0.1180 | -0.1675 |
| | | (0.6551) | (6.2727) | (0.2937) | (0.3265) |
| | CA | **0.4786 | *0.3149 | 0.2778 | 0.1635 |
| | | (8.6737) | (4.9284) | (3.4610) | (0.6989) |
| | CH | *-1.8738 | -0.0975 | 0.2134 | -1.4347 |
| | | (3.9779) | (0.0597) | (0.3340) | (2.3360) |
| | DE | **0.4674 | -0.0603 | **0.3382 | 0.2165 |
| | | (10.5903) | (0.2265) | (6.6328) | (1.6931) |
| | EC | **-1.2208 | **1.3172 | **-1.5110 | *-0.6264 |
| | | (14.3634) | (59.0855) | (16.6149) | (4.2417) |
| | FI | -0.3667 | **-0.7282 | -0.5280 | 0.0951 |
| | | (1.6644) | (7.2333) | (3.4612) | (0.1142) |
| | HK | 0.5912 | 0.4525 | 0.6137 | 0.0527 |
| | | (2.8204) | (1.8084) | (2.8599) | (0.0129) |
| | IN | **0.7379 | 0.3194 | *0.4052 | **0.8442 |
| | | (16.2999) | (3.6919) | (5.0517) | (18.2447) |
| | NO | -0.1308 | -0.7292 | -0.7091 | -0.1071 |
| | | (0.1132) | (3.1039) | (2.6210) | (0.0573) |
| | PL | 0.0083 | -0.4430 | 0.1058 | -0.1489 |
| | | (0.0003) | (0.7962) | (0.0912) | (0.0698) |
| | RU | 0.2868 | -0.1649 | -0.0972 | 0.2966 |
| | | (2.4931) | (1.0026) | (0.2518) | (2.0558) |
| | TW | **0.9403 | 0.3138 | **0.6225 | *0.5802 |
| | | (22.2354) | (2.7166) | (8.7272) | (6.2740) |
| Role | Buy | **-0.4363 | **-0.7622 | **0.2743 | **-0.3632 |
| | | (66.6035) | (197.2782) | (26.2916) | (33.3078) |
| NSS Before | F | -0.0114 | 0.1428 | -0.0795 | -0.0458 |
| | | (0.0179) | (3.1475) | (0.8377) | (0.2242) |
| Nego Experience | | *-0.1057 | -0.0774 | **-0.1660 | -0.0489 |
| | | (4.7215) | (2.7704) | (10.2065) | (0.7328) |
| Weight Issues | | 0.0550 | *-0.1632 | 0.0047 | 0.0103 |
| | | (0.5423) | (5.1453) | (0.0036) | (0.0135) |
| Weight Options | | -0.0489 | 0.0780 | -0.0335 | -0.0410 |
| | | (0.4663) | (1.3034) | (0.1991) | (0.2349) |
| Understand Case | | **0.2895 | **0.2192 | **0.2909 | **0.3289 |
| | | (20.2296) | (12.4488) | (17.9294) | (19.4620) |
| Age | | *0.0159 | *0.0148 | *0.0159 | **0.0267 |
| | | (4.3400) | (3.9791) | (3.9270) | (10.3766) |
| $R^2$ | | 0.0669 | 0.1234 | 0.0360 | 0.0384 |
| max. rescaled $R^2$ | | 0.1090 | 0.1871 | 0.0629 | 0.0738 |

Table 6: Factors influencing monotonicity errors: parameter estimates and $\chi^2$ values

| Variable | | Price | Delivery | Payment | Returns |
|---|---|---|---|---|---|
| Gender | F | 1.1000 | 0.8910 | 0.9320 | 0.9950 |
| Country | AT | 0.6130 | 0.4910 | 0.9680 | 0.6680 |
| | CA | 1.2130 | 1.2300 | 1.1350 | 0.9300 |
| | CH | 0.1150 | 0.8140 | 1.0650 | 0.1880 |
| | DE | 1.2000 | 0.8450 | 1.2060 | 0.9810 |
| | EC | 0.2220 | 3.3500 | 0.1900 | 0.4220 |
| | FI | 0.5210 | 0.4330 | 0.5070 | 0.8690 |
| | HK | 1.3580 | 1.4110 | 1.5890 | 0.8330 |
| | IN | 1.5720 | 1.2350 | 1.2900 | 1.8370 |
| | NO | 0.6600 | 0.4330 | 0.4230 | 0.7100 |
| | PL | 0.7580 | 0.5760 | 0.9560 | 0.6810 |
| | RU | 1.0010 | 0.7610 | 0.7800 | 1.0630 |
| | TW | 1.9250 | 1.2280 | 1.6030 | 1.4110 |
| Role | Buy | 0.4180 | 0.2180 | 1.7310 | 0.4840 |
| NSS Before | F | 0.9780 | 1.3310 | 0.8530 | 0.9120 |
| Nego Experience | | 0.9000 | 0.9250 | 0.8470 | 0.9520 |
| Weight Issues | | 1.0570 | 0.8490 | 1.0050 | 1.0100 |
| Weight Options | | 0.9520 | 1.0810 | 0.9670 | 0.9600 |
| Understand Case | | 1.3360 | 1.2450 | 1.3380 | 1.3890 |
| Age | | 1.0160 | 1.0150 | 1.0160 | 1.0270 |

Table 7: Odds ratios for table 6

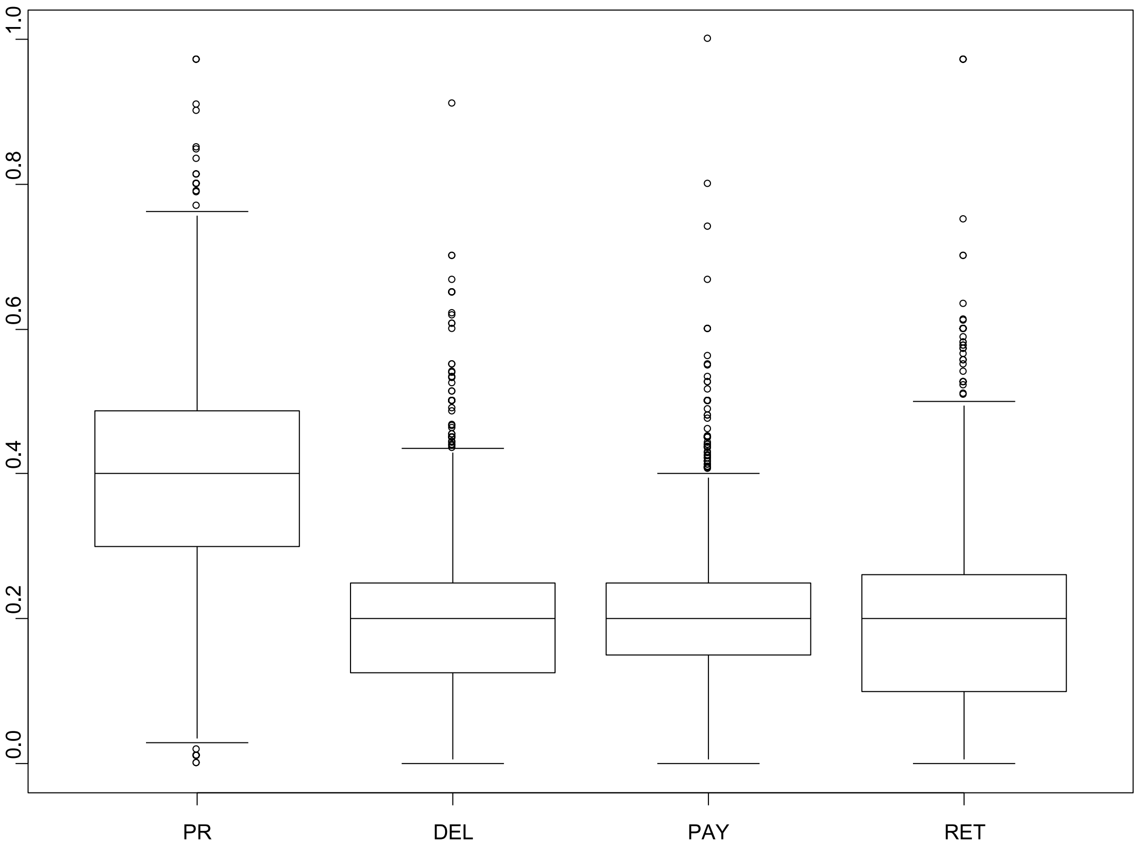


Figure 5: Distributions of weights for the four attributes price, delivery times, payment terms and returns, all users

The weights represent the importance of attributes to users. More precisely, they represent the importance users assign to the different attributes when acting their assigned role in the context of the case. Thus it is possible that assigning a different role can cause users to use different weighs for the attributes. This is indeed the case, as figure 6 shows.

While in both roles the highest weight was assigned to price, the average weights for price still are different between the two roles at the 0.01% level of significance as indicated by a Kruskal-Wallis test. Sellers on average used a weight which was about 3 percentage points higher than the weight used by buyers. Sellers also had a higher weight for the attribute terms of payment, while the weight for delivery times and returns policy was higher for buyers. Thus it seems that in the role of a seller, most users put more emphasis on attributes that relate to monetary aspects of the case, while for buyers, quality aspects

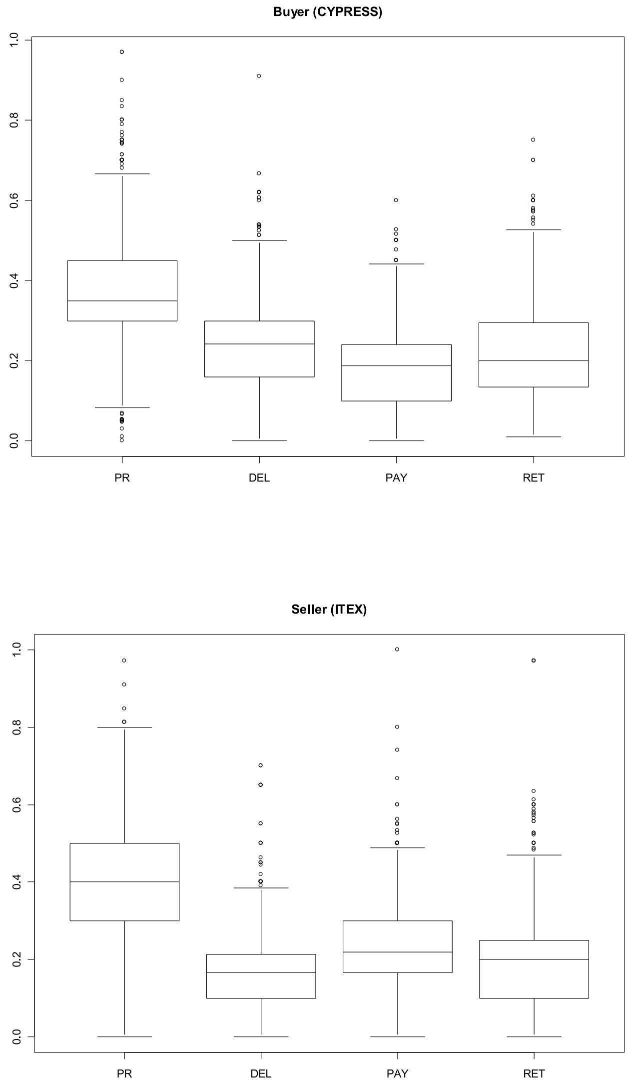


Figure 6: Distributions of weights for different roles

are more important. These differences in weights could be an alternative explanation for the differences in buyers' and sellers' valuations of goods which were identified in other empirical studies (Casey, 1995).

These differences are also reflected in the outcomes. While for price, the mean compromise value across all experiments was almost identical to the mean value of the five possible outcomes, the average compromise was considerable better for buyers than the mean attribute value for the attributes delivery time and returns, and better for the seller for the attribute payment. This indicates that subjects were indeed able to identify the potential for Pareto improvements that the different weights between the two roles imply.

| | | + Price | + Delivery | + Payment | + Returns |
|---|---|---|---|---|---|
| Role | Buyer | 0.3728 | 0.2347 | 0.1780 | 0.2144 |
| | Seller | 0.4030 | 0.1734 | 0.2308 | 0.1927 |
| | All | 0,3893 | 0,2012 | 0,2069 | 0,2026 |
| Average compromise | | 3.933 | *30.549 | *22.421 | *4.131 |
| Mean attribute value | | 3.930 | 38.750 | 30.000 | 5.000 |

+ significant difference in weight between buyers and sellers, $p < .0001$

* compromise significantly different from mean attribute value, $p < .0001$

Table 8: Average attribute weights by role and their effect on compromise values

Apart from the weights of individual attributes, the dispersion of weights across attributes also provides information about the preference structures of users. When a user differentiates strongly between the attributes, their weights should be more different than when a user considers all attributes to be similarly important. Table 9 shows the main results of a GLM model, in which the standard deviation of weights across attributes was regressed on various socio-demographic factors of users.

There is a strong effect of the role again, buyers on average use less different weights for the attributes than sellers. Users who found the weighting of issues more difficult also used more significantly more similar weights for attributes. On the other hand, older users differentiated more strongly between attributes. There are also some cultural effects. While in the standard design of GLM models, one country (in our case the US) is used as a reference category, and all other categories are compared to that reference category, we also compared the parameter estimate of each country against the average parameter estimate for all countries to obtain a clearer picture of countries exhibiting a different behavior. Using this method, we find that users from Austria, Switzerland, Russia, the US and (weakly significant) Hong Kong used more different weights, while users from Germany, Ecuador, Taiwan and (weakly significant) India used more equal weights than on average.

| Variable | | Estimate | t-value |
|---|---|---|---|
| Role | Buyer | **-0.0100 | (-4.9600) |
| Gender | F | -0.0017 | (-0.8000) |
| Country | AT | **0.0110 | (2.6500) |
| | CA | -0.0061 | (-1.9200) |
| | CH | **0.0192 | (3.2200) |
| | DE | **-0.0125 | (-4.8300) |
| | EC | **-0.0445 | (-10.4200) |
| | FI | -0.0079 | (-1.8500) |
| | HK | *0.0170 | (2.0500) |
| | IN | *-0.0071 | (-1.9900) |
| | NO | 0.0088 | (1.3200) |
| | PL | -0.0071 | (-1.0600) |
| | RU | **0.0366 | (12.3300) |
| | TW | **-0.0185 | (-4.5200) |
| | US | **0.0107 | (3.7400) |
| NSS Before | F | 0.0058 | (1.7300) |
| Nego Experience | | 0.0013 | (1.3600) |
| Weight Issues | | **-0.0053 | (-3.5300) |
| Weight Options | | -0.0014 | (-0.9400) |
| Understand Case | | -0.0010 | (-0.7400) |
| Age | | **0.0006 | (3.5400) |
| $R^2 = 0.0990$ | | | |

Table 9: Factors influencing weight dispersion

### 3.3 Concavity

In decision analysis and economic models, one typically assumes decreasing marginal benefits and consequently, the utility functions for each attribute should have a concave shape.

For our empirical analysis, this concept needs to be operationalized. A straightforward approach would classify a single attribute utility function as concave if it is concave for the entire range of attribute values and as convex if it is convex for the entire range. But this classification would imply a considerable loss of information, since single attribute utility functions might exhibit a strong or weak curvature.

In decisions under risk, a common measure for risk aversion, which is formally represented also by concavity of utility functions, is the Arrow-Pratt measure of risk aversion (Keeney

& Raiffa, 1976), which is defined as

$$A = -\frac{u''(x)}{u'(x)} \tag{3}$$

i.e. the coefficient of the second and first derivative of the utility function. However, it is only a local measure, while for our purpose we need a measure which describes the shape of the entire utility function.

Therefore, the following approach was taken: first the attribute ranges were all standardized to the [0,1] interval and the single attribute utility functions to $u(0) = 0$ and $u(1) = 1$ and decreasing functions were reversed to make them comparable to increasing functions. Then to each single attribute utility function, a negative exponential function of the form

$$u(x) = \frac{1 - e^{-\rho x}}{1 - e^{-\rho}} \tag{4}$$

was fitted using a least squares approach. Function (4) is convex for $\rho < 0$ and concave for $\rho > 0$ and approaches a linear function for $\rho \to 0$. Thus the parameter $\rho$ can be used as an indicator of the concavity of the utility function.

Table 10 shows the distribution of convex, linear and concave utility functions for the four attributes. For this table, a function was classified as linear when $|\rho| \leq 0.01$.

| | Shape | | |
|---|---|---|---|
| Attribute | convex | linear | concave |
| Price | 35.02 | 0.38 | 64.60 |
| Delivery | 42.19 | 4.07 | 53.75 |
| Payment | 21.77 | 23.86 | 54.37 |
| Returns | 29.27 | 23.38 | 47.35 |

Table 10: Distribution of different shapes of single attribute utility functions (in %)

Linear (or approximately linear) functions occur mainly for those attributes for which only three values are available. For Delivery, which had four possible values, their share drops to about 4%, for Price with five possible values it becomes negligible.

Only between half and two thirds of all subjects exhibited a concave single attribute utility function, while convex functions were rather frequent. Thus the usual assumption of decreasing marginal benefits is not reflected in the data to a large extent.

To analyze factors which influence the shape of single attribute utility functions, a regression analysis on parameter $\rho$ of the utility functions was performed. Table 11 shows the parameter estimates which were obtained from this analysis.

| Factor | | Price | Delivery | Payment | Return |
|---|---|---|---|---|---|
| Gender | F | -0.4731 | -1.7172 | -0.2543 | -0.2867 |
| | | (-1.75) | (-1.20) | (-1.21) | (-1.25) |
| Country | AT | ** -1.6971 | * -6.0279 | * -1.0163 | * -0.9059 |
| | | (-3.28) | (-2.20) | (-2.45) | (-2.02) |
| | CA | -0.4230 | -0.4641 | -0.1720 | -0.5716 |
| | | (-1.14) | (-0.23) | (-0.59) | (-1.82) |
| | CH | 0.1373 | 1.7039 | ** 2.6165 | -1.1264 |
| | | (0.16) | (0.35) | (3.60) | (-1.46) |
| | DE | ** -1.1481 | -1.9498 | ** -1.0784 | ** -0.7386 |
| | | (-3.44) | (-1.11) | (-4.16) | (-2.63) |
| | EC | ** 3.8925 | * 5.8243 | ** 2.8042 | ** 2.9891 |
| | | (7.88) | (2.12) | (7.09) | (6.86) |
| | FI | 1.0514 | * 6.4163 | 0.3587 | * 1.2057 |
| | | (1.92) | (2.22) | (0.82) | (2.51) |
| | HK | -0.7288 | 4.1647 | 0.7409 | 0.4762 |
| | | (-0.68) | (0.72) | (0.90) | (0.54) |
| | IN | 0.1044 | -3.5277 | * -0.7613 | 0.1446 |
| | | (0.22) | (-1.42) | (-2.11) | (0.36) |
| | NO | -0.2495 | -7.6905 | ** -1.6405 | -0.7388 |
| | | (-0.32) | (-1.87) | (-2.68) | (-1.09) |
| | PL | 0.4891 | 8.3186 | -0.1346 | -1.4417 |
| | | (0.56) | (1.80) | (-0.19) | (-1.89) |
| | RU | 0.5010 | 2.1381 | -0.3924 | 0.1886 |
| | | (1.19) | (0.96) | (-1.20) | (0.52) |
| | TW | -0.1462 | -3.3857 | 0.0123 | *1.2405 |
| | | (-0.23) | (-1.06) | (0.03) | (2.49) |
| | US | ** -1.7830 | ** -5.5201 | ** -1.3371 | * -0.7218 |
| | | (-4.94) | (-2.84) | (-4.74) | (-2.35) |
| Role | Buyer | ** -0.9725 | ** 13.4900 | 0.1213 | ** -1.5933 |
| | | (-3.72) | (-9.65) | (0.59) | (-7.13) |
| NSS Before | F | -0.1728 | 0.5013 | -0.0447 | -0.6602 |
| | | (-0.41) | (0.22) | (-0.13) | (-1.81) |
| Nego. Experience | | 0.0145 | *1.6845 | -0.0840 | -0.0879 |
| | | (0.11) | (2.41) | (-0.83) | (-0.80) |
| Weight Issues | | -0.2756 | ** -2.8009 | -0.0513 | -0.3075 |
| | | (-1.46) | (-2.77) | (-0.34) | (-1.89) |
| Weight Options | | * 0.4477 | * 2.3016 | 0.2621 | ** 0.4451 |
| | | (2.47) | (2.37) | (1.82) | (2.84) |
| Understand Case | | ** -0.5768 | * -2.1717 | ** -0.5674 | ** -0.4502 |
| | | (-3.30) | (-2.37) | (-4.18) | (-3.05) |
| Age | | ** -0.0606 | -0.1407 | -0.0107 | ** -0.0607 |
| | | (-2.75) | (-1.18) | (-0.61) | (-3.23) |
| $R^2$ | | 0.0660 | 0.0599 | 0.0493 | 0.0643 |

Table 11: Factors influencing the curvature of single attribute utility functions

There is a consistent influence of culture, users from Austria and the US exhibit less concave functions in all attributes, while users from Ecuador typically have more concave functions. Also users who understand the case less well have less concave utility functions, while users who find the weighting of options more difficult tend to have a more concave utility function. While it is tempting to relate these results to empirical studies on risk attitudes in different cultures (E. Weber & Hsee, 1998), it should be noted that utility functions where elicited in a risk-free setting and thus do not represent the users' risk attitudes. However, the fit of the models in general was rather weak, so the shape of single attribute utility functions seems to be mostly determined by factors not included in our data.

## 4 Consistency with Observed Behavior

While Inspire provides the users with considerable information based on their elicited utility functions, the system does not control their behavior during the negotiations. Thus it is possible that the actual behavior of users is inconsistent with the information encoded in their utility functions.

There are two types of decisions which users must make during a negotiation:

- Whether to accept or reject an offer made by their opponent and
- which proposals to make themselves.

When we consider the set of offers made by the opponent as the set of alternatives which is available to a negotiator, a rational negotiator should select the offer with the highest utility. Thus a violation of consistency occurs if the compromise to which the user has agreed does not have the highest utility among all offers made by the opponent.

One could argue that in a dynamic perspective, a negotiator might reject an offer because he expects to be able to extract more concessions from the opponent, and later on this turns out to be not possible. However, even in this setting it will usually be possible to revert to an offer which has previously been on the table, so settling for a compromise which is rated worse than a previous offer from the opponent can be considered as a violation of consistency. This view is also supported by the fact that Inspire negotiations take a rather short time of three weeks, and there are no external factors which would make returning to a previous offer impossible. This distinguishes our experiments from actual buyer/seller negotiations, where for example increased costs of the supplier would make it impossible to return to a previous offer containing a lower price.

Thus we consider a negotiator to violate consistency with respect to the opponent's offers when according to the negotiator's utility function, a rejected offer from the opponent should be preferred to the final compromise.

A similar argument can also be formulated for offers made by the negotiator. Assuming that the negotiation process consists of a sequence of concessions, each offer made by a negotiator should have a lower utility – from that negotiator's point of view – than the previous offers, and consequently the final compromise should have the lowest utility. While this view ignores the possibility that negotiators discover possibilities for joint gains during the bargaining process, it is nevertheless a quite common approach to negotiations and also compatible with dynamic bargaining models (Cross, 1965; Contini & Zionts, 1968; Bronisz, Krus, & Wierzbicki, 1988). We therefore consider a negotiator to violate consistency with respect to the negotiator's own offers if the utility of the final compromise exceeds the utility of an offer the negotiator has made before during the negotiation.

| | | Consistent with respect to | | |
|---|---|---|---|---|
| | | Opponent's | Own | |
| Role | Neither | Offers | Offers | Both |
| Buyer | 1.59 | 12.30 | 10.18 | 75.93 |
| Seller | 1.73 | 17.43 | 10.31 | 70.53 |
| All | 1.67 | 15.06 | 10.25 | 73.03 |

Table 12: Consistent and inconsistent behavior of negotiators (in %)

Table 12 gives an overview of the frequency of consistency violations. While most of the users behaved consistently, a violation of consistency occurred in about one quarter of all cases. In interpreting this number, it should be noted that a negotiator was classified as inconsistent if just one out of possibly several offers violated the conditions formulated above.

Violations of consistency are more frequent with respect to the negotiator's own offers than with respect to the opponent's offers. This might be related to the fact that the theoretical argument why one should be consistent with respect to the opponent's offer is much stronger. Violations of consistency with respect to one's own offers could still be considered as rational behavior if possibilities for joint gains are only gradually discovered during the negotiations. Surprisingly, inconsistencies with respect to the opponent's offer are more influenced by role effects than inconsistency with respect to one's own offers.

Apart from the role, another factor influencing inconsistent behavior could be the complexity of the preferences. It is plausible to assume that negotiators with more complex preferences more often behave inconsistently than negotiators with simple preferences. In terms of the utility function, complexity of preferences depends on several factors like the nonlinearity of the single attribute utility functions or the similarity/dissimilarity of weights.

To test this hypothesis, a logistic regression analysis was performed using the different forms of violation of consistency as dependent variable. Along with the complexity of preferences, socio-demographic factors were also considered as explanatory variables. The results of these analyses are summarized in table 13. This table presents both the maximum likelihood parameter estimates and the odds ratio, as well as the values of the $\chi^2$ test for the estimates.

The hypothesis that the probability of inconsistency is influenced by the complexity of the utility function is confirmed by this analysis, especially with respect to the negotiator's own offers. One is more likely to behave inconsistently and rank the compromise higher than at least one own offer, if the number of non-monotonic single attribute utility functions is higher, the number of linear single attribute utility functions is lower or the standard deviation of weights is higher. In addition, users who performed only one estimation of the utility function are only half as likely to make an error concerning their own offers than users who performed several estimations of the utility function, which can also be considered as an indicator of (subjective) complexity of the utility function. Similarly, users who understood the case less well and users who are less experienced in negotiations are also more likely to make errors.

# 5 Conclusions and Future Research

Our exploratory study has led to several insights concerning the ability of users to represent their preferences in the form of utility functions, and on the relationship between preferences expressed in this way and actual behavior.

One important conclusion which can be drawn from our result is that the difficulties of encoding preferences in the form of utility functions seem to be not as large as they are sometimes seen in the literature (Pomerol & Barba-Romero, 2000). Even users who made errors in understanding the case description were able to represent their preferences in the form of utility functions.

Several structural properties of the users' utility functions are clearly reflected in their behavior and the outcome of the negotiations: The monotonicity of the single attribute functions in most cases seems to correctly represent the direction into which a user wants to influence an attribute, even when this direction does not conform to the requirements of the case description. The importance of attributes to users as represented in the attribute weights is also clearly recognizable in the negotiation outcomes.

| | | Type of inconsistency | | | | | |
|---|---|---|---|---|---|---|---|
| | | Own | | Opponent | | Any | |
| | | Estimate | Odds | Estimate | Odds | Estimate | Odds |
| N. Offers | | **0.2606 | 1.2980 | **0.3999 | 1.4920 | **0.2123 | 1.2370 |
| | | (79.6073) | | (153.5197) | | (199.9253) | |
| Gender | F | 0.0240 | 1.0490 | 0.0244 | 1.0500 | -0.0311 | 0.9400 |
| | | (0.1885) | | (0.1403) | | (0.4127) | |
| Country | AT | -0.0913 | 0.6740 | -0.0226 | 0.6530 | -0.0367 | 0.7540 |
| | | (0.1524) | | (0.0060) | | (0.0346) | |
| | CA | * 0.3646 | 1.0630 | 0.1545 | 0.7790 | *0.3013 | 1.0570 |
| | | (5.4611) | | (0.6087) | | (4.8914) | |
| | CH | -0.8797 | 0.3060 | -1.4459 | 0.1570 | **-1.1300 | 0.2530 |
| | | (3.7457) | | (2.3610) | | (7.3198) | |
| | DE | 0.2038 | 0.9050 | **0.5606 | 1.1690 | **0.3306 | 1.0880 |
| | | (1.9613) | | (9.8289) | | (7.0690) | |
| | EC | -0.1418 | 0.6410 | **-1.8144 | 0.1090 | **-0.7580 | 0.3660 |
| | | (0.3350) | | (12.6861) | | (11.0083) | |
| | FI | -0.0055 | 0.7340 | -0.1943 | 0.5500 | -0.1796 | 0.6530 |
| | | (0.0006) | | (0.4945) | | (0.8964) | |
| | HK | **-1.6588 | 0.1410 | *0.8414 | 1.5480 | -0.5829 | 0.4360 |
| | | (7.6187) | | (4.0790) | | (2.2587) | |
| | IN | -0.2153 | 0.5950 | 0.2359 | 0.8450 | -0.2634 | 0.6010 |
| | | (0.8831) | | (0.7524) | | (1.7316) | |
| | NO | **1.0541 | 2.1190 | 0.6945 | 1.3370 | **1.0535 | 2.2420 |
| | | (10.7623) | | (3.1729) | | (13.7095) | |
| | PL | *0.7019 | 1.4900 | 0.0750 | 0.7190 | 0.4379 | 1.2110 |
| | | (4.9634) | | (0.0271) | | (2.2874) | |
| | RU | **-0.5868 | 0.4110 | 0.2731 | 0.8770 | *-0.3250 | 0.5650 |
| | | (10.6559) | | (1.7696) | | (4.9363) | |
| | TW | **0.9517 | 1.9130 | 0.2380 | 0.8470 | **0.9060 | 1.9350 |
| | | (29.2507) | | (0.9292) | | (30.6382) | |
| Role | Buy | -0.0774 | 0.8570 | 0.1077 | 1.2400 | 0.0059 | 1.0120 |
| | | (1.9584) | | (2.7800) | | (0.0148) | |
| NSS Before | F | -0.0823 | 0.8480 | -0.0602 | 0.8870 | -0.1185 | 0.7890 |
| | | (0.8998) | | (0.3276) | | (2.4014) | |
| Nego Experience | | **-0.1643 | 0.8490 | -0.0809 | 0.9220 | -0.0820 | 0.9210 |
| | | (10.3544) | | (1.6800) | | (3.2264) | |
| Weight Issues | | -0.0573 | 0.9440 | -0.0174 | 0.9830 | -0.1128 | 0.8930 |
| | | (0.4549) | | (0.0311) | | (2.3973) | |
| Weight Options | | -0.0735 | 0.9290 | -0.0520 | 0.9490 | -0.0426 | 0.9580 |
| | | (0.7856) | | (0.2958) | | (0.3618) | |
| Understand Case | | *0.1494 | 1.1610 | *0.2084 | 1.2320 | **0.2212 | 1.2480 |
| | | (4.5070) | | (6.5490) | | (12.6277) | |
| Age | | 0.0039 | 1.0040 | * 0.0239 | 1.0240 | * 0.0180 | 1.0180 |
| | | (0.1545) | | (4.8081) | | (4.3022) | |
| N Non-monotonic | | **0.7261 | 2.0670 | **0.1656 | 1.1800 | **0.6430 | 1.9020 |
| | | (211.8444) | | (7.0192) | | (184.7334) | |
| N Linear | | **-0.2965 | 0.7430 | 0.0658 | 1.0680 | -0.1238 | 0.8840 |
| | | (10.2428) | | (0.5032) | | (2.8140) | |
| One estimate | | **-0.6786 | 0.5070 | 0.2336 | 1.2630 | **-0.3034 | 0.7380 |
| | | (36.9306) | | (3.2495) | | (9.8454) | |
| SD Weights | | **2.1059 | 8.2150 | 0.0383 | 1.0390 | **1.8742 | 6.5150 |
| | | (6.9597) | | (0.0016) | | (6.9579) | |
| $R^2$/max. rescaled $R^2$ | | 0.1570 | 0.2481 | 0.0738 | 0.1450 | 0.1730 | 0.2464 |

Table 13: Factors influencing consistency

These two properties mainly concern preferences at an ordinal scale, users are well able to indicate that they prefer one level of an attribute over another level, or that they consider one attribute as more important than another attribute. But representing preferences on a metric scale seems to be more difficult. The common assumption of decreasing marginal benefits, which would require comparison of the differences between attribute values, is violated in a relatively large number of the single attribute utility functions. Consequently, we have found a considerable number of inconsistencies between actual behavior during negotiations and the utility values. This discrepancy between utility functions and behavior is even more surprising in view of the fact that the utility values of offers are displayed to users during all their interactions with the system. Thus users who violated consistency with respect to an offer from the opponent knowingly agreed to a compromise with a lower utility value than a contract which was offered them before.

Our results indicate that complexity of preferences might be an important factor in such paradoxical decisions, and it seems that complexity is not resolved for users by providing them with utility values. This insight could have consequences for the development of multi-attribute decision support systems, where additional features to help users to understand and handle complex preferences might be required.

This study has also shown that users' preferences are not only complex, but also more diverse than is commonly assumed and that for example convex or non-monotonic single attribute utility functions occur quite frequently. This result is also important for the development of multi-attribute decision support methods. Frequently, simulation studies are used to compare different methods, e.g. (Triantaphyllou & Mann, 1989; Fry et al., 1996; Stewart, 1993, 1996; Zanakis et al., 1998). In performing such studies, it is probably not sufficient to consider only standard specifications for utility functions which represent decreasing marginal utility, but a wider range of different utility functions should be used to generate realistic results.

Although our study is based on a large empirical database, it still has some limitations, which need to be addressed in future research. From an empirical point of view, the two drawbacks of the Inspire database are its limitation to a single negotiation case and the lack of control over the user population. While using a single case is an advantage for obtaining consistent data, it also limits the results to a specific context. Decision problems in other areas than buyer-supplier-negotiations might exhibit different empirical phenomena, which are not addressed in our study. The lack of control over the user population makes it difficult to estimate how representative the users of Inspire are for any more general population.

In addition, since the Inspire project has been started with different research questions in mind, and the questionnaires used in the system could not be changed to maintain consistency of data, not all the factors that might be important for the structure of utility functions could be analyzed in this study. This problem is clearly recognizable in the relatively low fit of some of the statistical models that were presented here.

Nevertheless, this study has helped to uncover several important relationships between preference structures and actual behavior of negotiators, which should be explored further. While we have already noticed that the importance of attributes, which is reflected in their weights, has a strong influence on the outcomes of negotiations, other factors encoded in the utility functions might also have similar effects. For example, a convex utility function means that a user has increasing marginal utility from an attribute. This could lead to a particularly tough bargaining behavior with respect to such attributes, since users would loose much by deviating from their optimal values. Future research will also address the issue of actual behavior in more detail, for example by looking into concession patterns of users and relating them to structural properties of utility functions. This research could, beyond the domain of negotiations, lead to general insights into the relationship between preference structure and behavior.

## References

Beroggi, G. E. (2000). An experimental investigation of preference elicitation methods in policy decision-making. *Journal of Multi-Criteria Decision Analysis*, *9*(1-3), 76-89.

Borcherding, K., Eppel, T., & Winterfeldt, D. von. (1991). Comparison of weighting judgements in multiattribute utility measurement. *Management Science*, *37*(12), 1603-1619.

Bronisz, P., Krus, L., & Wierzbicki, A. P. (1988). Towards interactive solutions in a bargaining problem. In A. Lewandowski & A. Wierzbicki (Eds.), *Aspiration based decision support systems* (p. 251-268). Berlin: Springer.

Buchanan, J. T., & Corner, J. L. (1997). The effects of anchoring in interactive multiple criteria decision making solution methods. *Computers and Operation Research*, *24*(10), 907-918.

Casey, J. T. (1995). Predicting buyer-seller pricing disparities. *Management Science*, *41*(6), 979-999.

Contini, B., & Zionts, S. (1968). Restricted bargaining for organizations with multiple objectives. *Econometrica*, *36*, 397-414.

Corner, J. L., & Buchanan, J. T. (1995). Experimental consideration of preference in decision making under certainty. *Journal of Multi-Criteria Decision Analysis*, *4*(2), 107-121.

Cross, J. C. (1965). A theory of the bargaining process. *American Economic Review*, *40*, 67-94.

Delquie, P. (1993). Inconsistent trade-offs between attributes: New evidence in preference assessment biases. *Management Science*, *39*(11), 1382-1395.

Delquie, P. (1997). 'Bi-matching': A new preference assessment method to reduce compatibility effects. *Management Science*, *43*(5), 640-658.

Fischer, G., Carmon, Z., Ariely, D., & Zauberman, G. (1999). Goal-based construction of preferences: Task goals and the prominence effect. *Management Science*, *45*(8), 1057-1075.

Fry, P., Rinks, D. B., & Ringuest, J. L. (1996). Comparing the predictive validity of alternatively assessed multi-attribute preference models when relevant decision attributes are missing. *European Journal of Operational Research*, *94*, 599-609.

Hadley, C. F., Schoner, B., & Wedley, W. C. (1997). A field experiment comparing anchored and unanchored criteria weights in the analytic hierarchy process. *Journal of Multi-Criteria Decision Analysis*, *6*(3), 140-149.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: J. Wiley & Sons.

Kersten, G., & Noronha, S. (1999). WWW-based negotiation support: Design, implementation, and use. *Decision Support Systems*, *25*(2), 135-154.

Kimbrough, S., & Weber, M. (1994). An empirical comparison of utility assesment programs. *European Journal of Operational Research*, *75*, 617-633.

Nitzsch, R. von, & Weber, M. (1993). The effect of attribute ranges on weights in multiattribute utility measurements. *Management Science*, *39*(8), 937-943.

Pomerol, J.-C., & Barba-Romero, S. (2000). *Multicriterion decision in management: Principles and practice*. Dordrecht: Kluwer.

Schoemaker, P. J. H., & Waid, C. C. (1982). An experimental comparison of different approaches to determining weights in additive utility models. *Management Science*, *28*, 182-196.

Stewart, T. (1993). Use of piecewise linear value functions in interactive multicriteria decision support: A Monte Carlo study. *Management Science*, *39*(11), 1369-1381.

Stewart, T. (1996). Robustness of additive value function methods in MCDM. *Journal of Multi-Criteria Decision Analysis*, *5*(4), 301-309.

Triantaphyllou, E., & Mann, S. (1989). An examination of the effectiveness of multi-dimensional decision-making methods: A decision-making paradox. *Decision Support Systems*, *5*(3), 303-312.

Wang, M., & Yang, J. (1998). A multi-criterion experimental comparison of three multi-attribute weight measurement methods. *Journal of Multi-Criteria Decision Analysis*, *7*(6), 340-350.

Weber, E., & Hsee, C. (1998). Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. *Management Science*, *44*(9), 1205-1217.

Weber, M., Eisenführ, F., & Winterfeldt, D. von. (1988). The effects of splitting attributes on weights in multiattribute utility measurement. *Management Science*, *34*, 431-445.

Zanakis, S. H., Solomon, A., Wishart, N., & Dublish, S. (1998). Multi-attribute decision making: A simulation comparison of select methods. *European Journal of Operational Research*, *107*(3), 507-529.