DRAFT – please do not quote

Framing and training to reduce starting point bias in choice experiments

Rolf A. Groeneveld

Tilburg University
Wageningen University

Contact address:
Wageningen University
PO Box 8130
NL-6700 EW Wageningen
The Netherlands
rolf.groeneveld@wur.nl

Abstract
Starting point bias is common problem in closed-ended non-market valuation studies. This paper analyses two possible methods to reduce starting point bias in choice experiments: letting respondents discover their preferences through a payment ladder; and framing the choice experiment in a familiar context. The two methods are applied in a valuation study of sustainable flatfish fisheries in the North Sea. Neither of the two methods have a significant impact on starting point bias, although the payment ladder treatment has a strong impact on respondents' sensitivity to the financial attribute. Moreover, respondents' WTP in the payment ladder correlates strongly with their WTP in the choice experiment.

Keywords: choice experiment, starting point bias, fisheries

1 Introduction

Closed-ended question formats in survey-based non-market valuation studies (e.g. dichotomous choice contingent valuation, conjoint analysis, and choice experiments) are generally regarded as easier to complete, and less prone to all sorts of bias, than open-ended formats (Whitehead, 2006). Despite these advantages, however, closed-ended question formats are inherently prone to starting point bias (SPB): the initial suggestion, be it the first question in a choice experiment or the first bid in a double bounded dichotomous choice CVM question, influences the respondents' eventual willingness to pay (WTP) or willingness to accept (WTA). After all, all such question formats necessarily start with a suggestion by the researcher to the respondent.

Most of the literature on SPB has so far focused on contingent valuation (e.g. Silberman and Klock, 1989; Mitchell and Carson, 1989; Ariely et al., 2003; Chien et al., 2005; van Exel et al., 2006; Bateman et al., 2007; Alevy et al., 2007). Ladenburg and Olsen (2006), investigates SPB in choice experiments. Generally, SPB appears to be stronger in surveys where the hypothetical market situation is difficult to understand, limited information is available on the public good, respondents are poorly informed on their preferences, and the starting point comes across as a plausible indicator of the 'true' value of the public good (van Exel et al., 2006). Contrary to the last argument, however, Ariely et al. (2003) shows that SPB even occurs when it is perfectly clear to the respondent that the starting point is as arbitrary as the respondent's social security number.

No silver bullet has yet been found to reduce starting point bias, but some suggestions have been made. Bateman et al. (2007) suggest letting respondents answer practice questions on a related subject before asking them to value the public good that is the purpose of the survey. One cannot rule out, however, that such exercises will influence the respondent's valuations. In their analysis of SPB in CEs, Ladenburg and Olsen (2006) suggest using different versions of the survey, each version offering a different starting point. Although this method gives insight into the effect of the starting point on the WTP estimate, it offers no clue of the 'right' estimate. Instead of training respondents, one could also consider putting the valuation questions in a familiar frame.

This paper tests two possible methods to reduce starting point bias in choice experiments. As has been argued in the literature before, uncertainty about one's own preferences and task complexity are possible sources of SPB, and the methods tested aim to tackle these two cources. The first method aims to 'train' respondents, i.e. to let them learn about their preferences by completing a payment ladder. Because a payment ladder offers a range of values rather than a point value, it may enable respondents to make up their own minds while influencing their findings only to a limited extent. The second method aims to 'frame' the choice experiment as a vote in an election, which is probably more familiar and less abstract than the choice between scenarios generally applied in choice experiments. This suggestion is based on Cosmides' (1989) finding that people have more difficulty solving an abstract task than one that is framed as a situation that is more familiar to them, even if both tasks are logically equivalent.

2 Experimental design

The choice experiment concerned respondents' preferences with respect to three attributes of sustainable fisheries and a change in their annual expenses as a fourth, financial, attribute (Table 1). The first fisheries attribute was the size of the North Sea flatfish stock (plaice and sole), that was either at the current level or at the level necessary to obtain Maximum Sustainable Yield. The second fisheries attribute was the damage caused to the North Sea benthos ecosystem by fishing activities. The beam trawl, which is the most common type of fishing gear in the Dutch flatfish fleet, is estimated to kill 30%-50% of living organisms on the North Sea bottom, whereas alternative gear types may be able to reduce these figures to 0-5%. The third fisheries attribute was the number of layover days in the plaice spawning period. During the spawning period (roughly from January to March), Dutch fishermen reduce fishing pressure on plaice to enable the species te reproduce, and also because the quality of plaice is significantly lower as the fish is quite thin and contains a lot of roe or sperm. Currently every vessel remains ashore for three weeks during this period, and in the CE it is proposed to extend this time to six weeks.

Table 1: Experimental design

Attribute	Flatfish stocks	Benthos mortality	Layover time in spawning period	Annual expenses
Level 1	Current level (Plaice 200 mln kg; Sole 24 mln kg)	Current (30%-50%)	Current (3 weeks)	+€10
Level 2	MSY level (Plaice 350 mln kg; Sole 50 mln kg)	0-5%	6 weeks	+€20
Level 3	-	-	-	+€30
Level 4	-	-	-	+€40

The respondents are participants in the CentER panel, which consists of more than 2000 households in the Netherlands. Panel members are selected to be representative of the Dutch population and complete online surveys on a regular basis. The sample is divided in 2^3 treatment groups along the distinctions ‘high starting point - low starting point’, ‘payment ladder - no payment ladder’, and ‘scenarios – elections’ (Table 2). Each respondent was assigned randomly to one of the eight treatment groups when he or she logged in to the online survey.

Table 2: Sample size per treatment group

Treatments		No ladder		Ladder		Total
		Elections	Scenarios	Elections	Scenarios	
Starting point	Low	231	242	226	239	938
	High	243	238	211	242	934
Total		474	480	437	481	1872

The survey explained each attribute one by one in 150 words maximum. After each explanation, the respondents in the ladder treatment groups completed a payment ladder as given in Table 3. The payment ladder question was accompanied by a small text asking respondents explicitly to carefully consider whether they really would be willing to pay this amount.

Table 3: Example of a payment ladder question

Are you willing to pay the following payments for an increase of the flatfish stocks to MSY level ? Please tick one answer in each row:			
	Absolutely	Maybe	Absolutely not
€ 0	o	o	o
€ 5 per year	o	o	o
€ 10 per year	o	o	o
€ 20 per year	o	o	o
€ 35 per year	o	o	o
€ 50 per year	o	o	o
€ 75 per year	o	o	o
€ 100 per year	o	o	o

After the explanations, respondents were shown an example of a CE question (Appendix A). To the election treatment groups this example and the actual CE questions were framed as choice between political parties, whereas to the scenario treatment groups the questions were framed as choices between policy scenarios. For the low starting point treatment groups the changes in annual expenses were €10 per year as shown in Appendix A, whereas for the high starting point groups the change in annual expenses were €40 per year.

Each CE question consisted of three options, namely the status quo with all fisheries attributes at their current level and the financial attribute at zero, and two alternative options. The eight alternative options are composed according to a orthogonal fractional factorial design. The CE also included a small text asking respondents explicitly to carefully consider whether they really would be willing to pay this amount.

3 Estimation

The survey data were analyzed by estimating the following utility function (suppressing case-specific indices):

$$U_i = a_i + \sum_j b_j X_{ij} + \varepsilon_i, \qquad (1)$$

where U_i is the utility derived from alternative i, a_i is an alternative-specific constant, b_j is the utility function's coefficient for attribute j, X_{ij} is the score of alternative i on attribute j and ε_i is a random component.

Model choice depends greatly on one's assumption on the probability distribution of ε_i. If one can assume that ε_i is identically distributed for all i's, and that all error terms are independent, one can estimate equation (1) by conditional logit (Greene, 2003). The multinomial probit model allows for a more general specification of the covariances between the error terms, but is more computationally demanding. The IIA assumption was tested by estimating a multinomial probit model with and without restrictions on the covariance matrix (cf. Long and Freese, 2006). These tests strongly rejected the assumptions of identically and independently distributed error terms ($P<10^{-5}$), although the independence assumption had a P-value of .019 when interactions were included between the ladder treatment and the financial attribute. It was therefore decided to proceed with the multinomial probit model, and to estimate it with the ASMPROBIT function in Stata (Long and Freese, 2006).

Because U_i is essentially an ordinal measure without a natural scale or level[1], the covariance matrix is not identified unless particular constraints are imposed (Train, 2003). The ASMPROBIT function does so by taking the difference between each alternative's utility and the utility of the base alternative, and fixing the variance of a second alternative to 2. In model (1), this implies that the 3x3 covariance matrix is reduced to a 2x2 matrix. Because these matrices are symmetric, the model only estimates one variance and one covariance. In this paper the status quo alternative (alternative 1) was chosen as the base alternative, and one of the other alternatives was chosen as the scale alternative (alternative 2). Hence, the model estimates the variance of the difference between error terms ε_1 and ε_3, and the covariance of the difference between error terms ε_1 and ε_2 and that between error terms ε_1 and ε_3.

4 Results

Appendix B shows the estimates for five different multinomial probit models. The results suggest a strong influence of the interaction between a ladder treatment dummy and the financial attribute. If this interaction is not included (models (1) – (3)), all coefficients are significant, but the correlation coefficient as estimated from the covariance matrix is very close to 1. Although some correlation between alternatives may occur in reality, this extreme value indicates that some variation is not accounted for in the model that should be included. Including the interaction between ladder treatment dummy and financial attribute reduces the correlation coefficient to a value that is still high, but significantly different from 1. This interaction is therefore included in all subsequent estimations.

The strong impact on the estimates of the interaction term between ladder treatment and financial attribute also suggests that the ladder treatment has a strong impact on respondents' behaviour. In models (4) and (5), the interaction term has taken all variation from the financial attribute itself, suggesting that respondents who completed a payment ladder are more sensitive to costs than respondents who did not.

[1] I.e., neither adding a constant to all *U*'s nor multiplying all *U*'s by a constant affects the respondent's choices.

The interaction between a dummy for starting point treatment group and the financial attribute is positive with P<0.001 (Apendix B). This suggests that respondents whose CE example question valued the alternative options at a €40 increase in annual expenses were less sensitive to the financial attribute than were those whose alternative options in the CE example cost only €10 per year. A likelihood ratio test (Table 4) also suggests a significant impact of the starting point treatment on the estimations.

Table 4: Result of likelihood ratio test of starting point bias

	Low SP group	High SP group	Sum	Pooled	LR	P(LR)
LL	-3040.35	-2948.95	5989.30	-6007	34.66	.0001
#	938	934	1872	1872		

4.1 Effect of ladder treatment on starting point bias

Table 5 shows for each bid the number of respondents who had this particular bid as his or her lower or upper bound, where the lower bound is defined as the largest amount the respondent is absolutely willing to pay, and the upper bound is defined as the smallest amount the respondent is absolutely *not* willing to pay. Note that for each fisheries attribute 90-100 respondents were not willing to pay €0. This may indicate that these respondents did not fully understand the question, as it seems unlikely that they were genuinely opposed to the measures proposed.

Table 5: Upper and lower bounds of respondents' willingness to pay

	Flatfish stocks		Benthos mortality		Layover period	
	Lo	Up	Lo	Up	Lo	Up
€ 0	273	90	337	89	364	109
€ 5	138	67	119	60	103	66
€ 10	192	133	179	128	189	129
€ 20	118	156	150	150	141	133
€ 35	66	155	70	143	67	147
€ 50	56	124	56	115	45	121
€ 75	2	62	7	38	9	33
€ 100	0	130	0	131	0	122
Average WTP	€ 12.11	€ 37.15	€ 12.16	€ 37.70	€ 11.43	€ 35.06

What is most striking about these figures is the similarity between the attributes: for each attribute, respondents' average lower bound is about €12 and their average upper bound is about €37. This contrasts with the CE estimates, which suggest substantial differences in WTP between the three attributes. After all, models (1)-(3) in Appendix B suggest that WTP for benthos mortality should be 1.5 to 2 times as large as that for flatfish stocks, and models (4)-(5), which are probably more reliable as they include the interaction term between the ladder treatment and the financial attribute, suggest it should be about 5 times as large.

We have already seen that including the interaction between the ladder treatment dummy and the financial attribute has a strong impact on the estimations: the correlation coefficient is much lower, the alternative-specific constants and the cost coefficient lose significance and the differences between the coefficients of the three fisheries attributes are larger. The significance of the interaction suggests that respondents in the ladder treatment groups have on average a lower WTP for any attribute than those in the groups without the ladder treatment. The results show no evidence, however, that the treatment reduces starting point bias. Model (6) in Appendix B includes two interaction terms between the starting point treatment dummy and the financial attribute: one for all groups and one specifically for the ladder treatment group. If the ladder treatment would reduce starting point bias, one would expect the group-specific interaction term to be negative and statistically significant. It is, however, not significantly different from zero. Estimation with the interaction term split between a ladder and a non-ladder term gives very similar estimates for both terms, which suggests that the impact of the starting point does not change with the ladder treatment. A Likelihood Ratio test even suggests that the no-ladder treatment group shows a smaller difference between the low starting point and the high starting point than the ladder treatment group (Table 6).

Table 6: Tests for ladder-treatment effects in starting point bias

		Low SP group	High SP group	Sum	Pooled	LR	P(LR)
No ladder	LL	-1432.22	-1405.25	-2837.47	-2845.66	16.37	.0374
	#	473	481	954	954		
Ladder	LL	-1601.10	-1533.39	-3134.49	-3146.07	23.16	.0032
	#	465	453	918	918		

So if the WTP as estimated by the CE shows such larger differences between the attributes than does the WTP estimated by the payment ladder, do they correlate at all? A probit model that includes interactions between the ladder WTP and the attributes suggests there is (Appendix C). If this WTP would be a perfect predictor of the outcomes in the CE the coefficients of the attributes should not be significant, which is not the case. The significance of the interaction term, however, suggests that ladder WTP is at least an imperfect predictor of CE WTP.

4.2 Effect of frame treatment on starting point bias

The effect of the frame treatment on starting bias is estimated the same way as that of the payment ladder treatment, with similar results. Model (7) in Appendix B shows that the treatment-specific interaction term is not significantly different from zero, suggesting that framing the questions as either an election or an abstract choice between scenarios has no influence on starting point bias. A Likelihood Ratio test suggests that the impact of the starting point may be smaller in the election treatment group than in the scenario treatment group, but the difference in LR is quite small.

Table 7: Tests for frame-treatment effects in starting point bias

		Low SP group	High SP group	Sum	Pooled	LR	P(LR)
Election	LL	-1508.13	-1442.61	- 2950.74	-2958.77	16.07	.0654
	#	457	454	911	911		
Scenarios	LL	-1522.94	-1500.84	- 3023.78	-3086.82	26.08	.0020
	#	481	480	961	961		

One could argue that the effect is not significant because many people do not actually compare programs when they vote: they simply vote for the same party every election. Hence, they have received insufficient ‘real life training’ to make the election frame effective. To investigate this explanation the CE survey also included a question how respondents make up their mind at elections, so it was possible to identify those who actively compare party programs. Adding another specific interaction term for this group, however, has no effect on the estimations: this interaction term is also not significant.

5 Discussion and conclusions

This paper tested two methods to reduce starting point bias in choice experiments. The first method aims to 'train' respondents, i.e. letting them complete a payment ladder to familiarize them with their own preferences before completing the actual choice experiment. The second method aims to 'frame' respondents, i.e. putting the choice experiment questions in a context that is familiar to respondents, and that they have dealt with in the past.

The 'training' method has a strong impact on respondents' answers, but not the intended one. First, respondents do become more sensitive to the cost aspect after they have completed the payment ladder. The payment ladder may have made respondents more aware of their own preferences, or respondents may have taken its range as a hint to what is reasonable. The evidence is unclear as to which explanation is the most plausible. On one hand, WTP estimated by the CE varies more strongly with the public goods considered than that estimated by the payment ladder. This suggests that respondents just choose an option in the payment ladder regardless of the public good considered. On the other hand, average WTP as estimated by the payment ladder is somewhere in the lower region of the ladder, rather than somewhere in the middle. Moreover, ladder WTP correlates strongly to the answers in the CE. It therefore seems likely that respondents do take some *a priori* preferences into account when completing the payment ladder, and that the payment ladder makes them more aware of it. The treatment has no measurable effect, however, on their susceptibility to starting point bias.

The 'framing' method has little or no impact on the estimations. One possible explanation for this is that the presumed 'real life training' takes place with insufficient frequency to really train respondents, but it is also possible that the two treatments are still too similar to make a difference.

References

Alevy, J. E., C. E. Landry and J. A. List. 2007. *Anchors away: A field experiment on anchoring of consumer valuations*. Greenville, NC: East Carolina University.

Ariely, D., G. Loewenstein and D. Prelec. 2003. ""Coherent Arbitrariness": Stable demand curves without stable preferences." *Quarterly Journal of Economics* 118(1):73-105.

Bateman, I. J., D. Burgess, B. Day, W. G. Hutchinson and D. I. Matthews. 2007. "Preference learning versus coherent arbitrariness: Contrasting NOAA guidelines with a Learning Design Contingent Valuation (LDCV)." Paper presented at the 15th Annual Conference of the European Association of Environmental and Resource Economists, Thessaloniki, 27-30 June 2007.

Chien, Y.-L., C. J. Huang and D. Shaw. 2005. "A general model of starting point bias in double-bounded dichotomous contingent valuation surveys." *Journal of Environmental Economics and Management* 50(2):362-77.

Cosmides, L. 1989. "The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task." *Cognition* 31(3):187-276.

Greene, W. H. 2003. *Econometric Analysis*. New Jersey: Prentice Hall.

Ladenburg, J. and S. B. Olsen. 2006. *Starting point anchoring effects in choice experiments*. Kopenhagen: Institute of Food and Resource Economics.

Long, J. S. and J. Freese. 2006. *Regression models for categorical dependent variables using Stata*. College Station, Texas: Stata Press.

Mitchell, R. C. and R. T. Carson. 1989. *Using surveys to value public goods: the contingent valuation method*. Washington, DC: Resources for the Future.

Silberman, J. and M. Klock. 1989. "The behavior of respondents in contingent valuation: Evidence on starting bids." *Journal of Behavioral Economics* 18(1):51-60.

Train, K. E. 2003. *Discrete choice models with simulation*. Cambridge, UK: Cambridge University Press.

van Exel, N. J. A., W. B. F. Brouwer, B. van den Berg and M. A. Koopmanschap. 2006. "With a little help from an anchor: Discussion and evidence of anchoring effects in contingent valuation." *Journal of Socio-Economics* 35(5):836-53.

Whitehead, J. C. 2006. "A practitioner's primer on the contingent valuation method." In *Handbook on Contingent Valuation*, ed. Alberini, A. and J. R. Kahn, pp. 66-91. Cheltenham, UK: Edward Elgar.

Appendix A Example choice experiment question shown to respondents in the low starting point, election frame treatment groups

In the following series of questions we will look at fisheries policy with all three policy measures discussed previously. We will ask you four times to choose between three policy scenarios, that differ with regard to the long-term flatfish stocks, the mortality of life at the North Sea bottom due to fishing pressure, the layover time in the plaice spawning period and a change in your annual expenses.

Elections

The choice between policy scenarios is comparable to a choice between political parties during elections. Therefore, try to imagine it is election time, and you can choose between three political parties. The only differences between these parties regard fisheries policy and your annual expenses. As regards other policy areas, such as defense and health care, you can assume that the parties propose the same policies. There will always be one party that aims to maintain the current situation. Which party would you vote for?

Below is an example of such a question.

Which political party would you vote for if the table below gives their political programs?

	Party 1 (current situation)	**Party 2**	**Party 3**
Flat fish stocks	Plaice 200 mln kg; Sole 24 mln kg	Plaice 200 mln kg; Sole 24 mln kg	Maximum Sustainable Yield, or Plaice 350 mln kg; Sole 50 mln kg
Benthos mortality due to fishing	30%-50%	0-5%	30%-50%
Layover time in spawning period	3 weeks	3 weeks	6 weeks
Annual expenses	No change	€10 more than now	€10 more than now
Choice: o I prefer scenario 1 o I prefer scenario 2 o I prefer scenario 3			

Party 1 versus 2

Perhaps you think reducing the mortality of North Sea bottom organisms is not worth paying €10 per year. In that case you will probably prefer Party 1 over Party 2.

Party 1 versus 3

Perhaps you are willing to pay at least €10 per year to increase flatfish stocks to MSY-level and reduce fishing pressure on plaice in the spawning period. In that case you will probably prefer Party 3 over Party 1.

Party 2 versus 3
Perhaps flatfish stocks are more important to you than North Sea bottom organisms. In that case Party 3 might appela more to you than Party 2.

As you see, in every question you will be asked to make a careful trade-off between flatfish stocks, the mortality of life on the bottom of the North Sea, the layover time in the plaice spawning period and you annual expenses.

Appendix B Probit estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ASC2	.9968***	1.156***	.9949***	-.1012	.1046	.1189	.1162	.0791
	(.1407)	(.1396)	(.1417)	(.2649)	(.2485)	(.2479)	(.2478)	(.2548)
ASC3	1.061***	1.179***	1.059***	.1818	.3518*	.3635	.3613	.3309
	(.1028)	(.1032)	(.1030)	(.2167)	(.2002)	(.1995)	(.1995)	(.2057)
Stocks	.1626***	.1522***	.1628***	.2233***	.2071***	.2062***	.2064***	.2089***
	(.0163)	(.0148)	(.0163)	(.0301)	(.0270)	(.0268)	(.0268)	(.0275)
Benthos	.2974***	.2364***	.2983***	.9078***	.8052***	.7977***	.7989***	.8184***
	(.0552)	(.0497)	(.0562)	(.1393)	(.1300)	(.1295)	(.1295)	(.1333)
Layover	.2402***	.2218***	.2406***	.3859***	.3668***	.3653***	.3656***	.3694***
	(.0272)	(.0217)	(.0278)	(.0404)	(.0398)	(.0397)	(.0397)	(.0401)
Costs	-.0139***	-.0169***	-.0137***	.0023	-.0013	-.0016	-.0018	-.0015
	(.0019)	(.0020)	(.0020)	(.0028)	(.0031)	(.0030)	(.0030)	(.0030)
Interactions with costs:								
Election	-	-	-.0002	-	-.0009	-	-	-
			(.0008)		(.0013)			
Ladder	-	-	-	-.0128***	-.0126***	-.0130***	-.0126***	-.0126***
				(.0014)	(.0014)	(.0019)	(.0014)	(.0014)
Interactions between costs and starting point for subgroups:								
All	-	.0038***	-	-	.0048***	.0043*	.0050**	.0050**
		(.0008)			(.0013)	(.0019)	(.0016)	(.0016)
Ladder	-	-	-	-	-	.0010	-	-
						(.0025)		
Elections	-	-	-	-	-	-	-.0004	-.0015
							(.0018)	(.0023)
Elections #)	-	-	-	-	-	-	-	.0020
								(.0027)
LL	-6053	-6042	-6053	-6007	-6000	-6000	-6000	-6000
lnl2_2	-2.9319	-3.8281	-2.8935	-.5028***	-.5951***	-.6032***	-.6026***	-.5814***
	(4.0248)	(15.093)	(3.7851)	(.1428)	(.1594)	(.1608)	(.1606)	(.1591)
l2_1	.9807***	1.050***	.9800***	.6460***	.7407***	.7465***	.7453***	.7297***
	(.0574)	(.0558)	(.0572)	(.1177)	(.1006)	(.1000)	(.0998)	(.1046)
ρ_{21} a)	.9985***	.9998***	.9984***	.7300***2)	0.8021***1)	.8066***1)	.8060***1)	.7938***1)
	(.0118)	(.0065)	(.012)	(.1043)	(.0789)	(.0775)	(.0777)	(.0832)

*:P<0.05; **:P<0.01; ***:P<0.001

a) Point estimate, standard error and P-value calculated by Delta Method

1) Different from 1 with P<0.05

2) Different from 1 with P<0.01

#) Only respondents who compare party programs during elections

Appendix C Tests for correlations in WTP between payment ladder and CE

	Lower bound	Average	Upper bound
ASC2	-1.1311*** (.1909)	-.9459*** (.2072)	-.7217** (.2250)
ASC3	-.6736*** (.1420)	-.5033** (.1577)	-.3079 (.1747)
Stocks	.1483* (.0597)	.0407 (.0636)	.0673 (.0580)
Interaction stocks and WTP according to payment ladder	.0155*** (.0021)	.0130*** (.0016)	.0075*** (.0010)
Benthos	1.1936*** (.1068)	1.0605*** (.1156)	.9993*** (.1216)
Interaction benthos and WTP according to payment ladder	.0179*** (.0018)	.0114*** (.0013)	.0054*** (.0008)
Layover	.2627*** (.0599)	.1583* (.0637)	.1902** (.0610)
Interaction layover period and WTP acc to payment ladder	.0180*** (.0025)	.0154*** (.0018)	.0091*** (.0011)
Costs	-.0062* (.0028)	-.0076** (.0029)	-.0086** (.0030)
Interaction starting point and cost attribute	.0035* (.0016)	.0037* (.0017)	.0040* (.0017)
LL	-2976	-2974	-3027
lnl2_2	-.2272* (.0907)	-.2371* (.0942)	-.3000** (.1085)
l2_1	-.0393 (.1169)	.1200 (.1193)	.2988* (.1207)
ρ_{21} a)	-.0493 1) (.1459)	.1503 1) (.1507)	.3741 *1) (.1485)

*:P<0.05; **:P<0.01; ***:P<0.001

a) Point estimate, standard error and P-value calculated by Delta Method

1) Different from 1 with $P<10^{-5}$