Asking for Individual or Household Willingness to Pay for Environmental Goods? Implication for aggregate welfare measures

Lindhjem, Henrik and Navrud, Ståle

Norwegian University of Life Sciences, Econ Pöyry

31 January 2008

Asking for Individual or Household Willingness to Pay for Environmental Goods? Implication for aggregate welfare measures

Henrik Lindhjem[a,b1] and Ståle Navrud[a],

[a] Department of Economics and Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway. henrik.lindhjem@umb.no; stale.navrud@econ.no

[b] ECON Pöyry, P.O. Box 5, N-0051, Oslo, Norway. www.econ.no

Revise and resubmit to *Environmental and Resource Economics* for its "Special Issue: The household, gender, children and environmental economics"

Presented at the 16th Annual Conference of the European Association of Environmental and Resource Economists, Gothenburg, Sweden, 25 - 28 June 2008.

[1] Corresponding author: henrik.lindhjem@umb.no

Asking for Individual or Household Willingness to Pay for Environmental Goods? Implication for aggregate welfare measures

Abstract

The aggregate welfare measure for a change in the provision of a public good derived from a contingent valuation (CV) survey will be much higher if the same elicited mean willingness to pay (WTP) is added up over individuals rather than households. A trivial fact, however, once respondents are part of multi-person households it becomes almost impossible to elicit an "uncontaminated" WTP measure that with some degree of confidence can be aggregated over one or the other response unit. The literature is mostly silent about which response unit to use in WTP questions and in some CV studies it is even unclear which type has actually been applied. We test for differences between individual and household WTP in a novel, web-administered, split-sample CV survey asking WTP for preserving biodiversity in old-growth coniferous forests in Norway. Two samples are asked both types of questions, but in reverse order, followed by a question with an item battery trying to reveal why WTP may differ. We find in a between-sample test that the WTP respondents state on behalf of their households is not significantly different from their individual WTP. However, within the same sample, household WTP is significantly higher than individual WTP; in particular if respondents are asked to state individual before household WTP. Our results suggest that using individual WTP as the response unit would overestimate aggregate WTP, and thus bias welfare estimates in benefit-cost analyses. Thus, the choice of response format needs to be explicitly and carefully addressed in CV questionnaire design in order to avoid the risk of unprofitable projects passing the benefit-cost test.

Keywords: Contingent valuation, household, individual, WTP; **JEL Classification**: Q51, H41

Introduction

The aggregate welfare measure for a change in the provision of a public good derived from a contingent valuation (CV) survey will be much higher if the same elicited mean willingness to pay (WTP) is added up over (adult) individuals rather than households. A trivial fact, however, once respondents are part of multi-person households, it becomes almost impossible to elicit an "uncontaminated" WTP measure that with some degree of confidence can be aggregated over one or the other response unit (e.g. Quiggin (1998), Bateman and Munro (2006)). The correct unit will not only depend on how and to whom the WTP question is phrased, but on the respondent's self-perceived agency and the type of resource allocation model prevailing in her[2] household (Delaney and O'Toole 2006; Strand 2007). Failing to appreciate this problem has important implications for the credibility of welfare estimates from CV studies, and stated preference research more generally (as for example acknowledged by Boyle (2003))[3]. The response unit issue has received little attention in the otherwise extensive CV literature, though resulting response unit biases in welfare estimates could be even higher than other more "high-profile" CV biases, summarised for example in Carson et al (2001). The result is an ambiguous and ill-informed CV practice applying a mix of approaches asking respondents for their personal WTP, their WTP on behalf of the household, or even leaving the unit unspecified in the WTP questions. No primary valuation studies we have seen, not even high-budget best practice studies such as Banzhaf et al (2006), investigate the sensitivity of welfare estimates to the implicit

[2] The gender-neutral way of varying the pronoun may be confusing in a paper on household preferences, where gender differences are well known. However, no such differences are intended unless explicitly stated.

[3] In fact, this is part of a wider problem as traditional microeconomics typically leave the definition of the "consumer" rather obscure – treating households and individual consumers the same (Vermeulen 2002).

assumptions about household decision-making underlying the choice of response unit[4]. Referring to CV case studies in the context of household decision-making models both in developing and developed countries, Smith and Van Houtven (2004: p164) state that: "*What is missing in all these cases is a systematic documentation of the consequences of using one set of maintained assumptions versus another to organize observed choices and estimate benefits people receive from improvements in environmental quality. The collective model offers one such perspective"*.

The aim of this paper is to investigate the empirical consequences of the choice of response unit – household or individual – and to inform the theoretical debate about household decision-making models in the context of CV of environmental goods. We attempt to answer the following questions: (1) What is the relationship between mean individual and mean household WTP?; (2) Do respondents within the same sample change their WTP when prompted instead to state individual or household WTP, and if so what are their stated reasons?; and finally; (3) Can household and respondent characteristics explain the observed relationship between household and individual WTP? The spill over and use of models from the large household decision-making literature to answer such questions has been limited to date within environmental valuation research. No generally agreed theoretical framework to analyse household decision-making in environmental valuation has been developed, though some attempts have been made (notably Quiggin (1998), Bergstrom (2003), Strand (2005; 2007), Smith and Van Houtven (1998; 2004), Munro (2005) and Bateman and Munro (2003)). We compare Strand's (2007) collective household decision model, which predicts that

[4] A few, limited benefit transfer studies have realised the potentially large implications for aggregation if WTP estimates are assumed to be for households or individuals.

response unit bias may actually not be a problem in large samples, with other models deriving hypotheses we test within and between samples. This paper is, to our knowledge, the first empirical contribution investigating the relationship between household and individual WTP for CV of environmental goods. Our study also supplements the results from a few limited empirical studies for other types of goods (e.g. TV broadcasting, health risks) or valuation methods (choice experiments) (Bateman and Munro 2006; Beharry and Scarpa 2006; Delaney and O'Toole 2006; Hasler et al. 2008; Delaney and O'Toole In press). Compared to previous research, we add several new and interesting dimensions: (1) Core elements of our CV questionnaire (e.g. type of good, scenario, payment vehicle, budget reminders) are consistent with eliciting both household and individual WTP from individual respondents; (2) Within this framework, all respondents are given both household and individual WTP questions, but the order is varied between samples, offering a clean test of differences in WTP from these two response units; finally (3) We utilise both respondents' stated reasons and a CV dataset supplemented with information from a pre-recruited web panel of respondents to explain the observed differences in individual and household WTP. We find that people state a higher WTP on behalf of the household than as individuals, though this difference is not significant at the 5 % level. However, when people are prompted to answer using the other response unit, the WTP difference increases, especially if they have been asked individual WTP first. Results suggest that response unit uncertainty may continue to be a source of substantial noise in aggregate welfare estimates, unless the issue is much more carefully addressed in survey design and testing.

Theoretical framework and empirical expectations

Theoretical framework

In their landmark book on CV, Mitchell and Carson (1989: p265-266) advise for pure public goods simply to "allow an adult who claims to be the household head" to state WTP on behalf of the household[5]. In a footnote they refer to Becker (1981) for implications of using household as the unit of analysis. Subsequent CV guidance documents have continued to treat lightly the appropriate response unit and potential implications of this choice (Arrow et al. 1993; Bateman et al. 2002; SEPA 2006). The literature that has developed in the area of family economics building on Nobel laureate Gary Becker's early contributions in the 1970s (Pollak 2003), give insights of potentially great significance also for environmental valuation research. This literature studies models of resource allocation in the family and how the effects of government policies depend on the dynamics of household decision-making (see e.g. Bergstrom (1994) and Vermeulen (2002) for reviews). The models are often divided into either unitary or collective. In the unitary model household choices can be described as if they were made by a single individual (Samuelson 1956; Becker 1973), the traditional assumption which greatly simplifies economic analysis of consumer behaviour,

[5] The reasons Mitchell and Carson put forward for this advice were in our view both misguided. First, they claimed that most payments for pure public goods are made at the household level, using income tax as an example. But income tax was always measured out and paid individually. Second, they claimed that choosing a household head was the U.S. Census Bureau practice at the time, though this rather archaic practice was abandoned already in 1980 partly due to feminist critique (Presser 1998). For quasi-private goods such as hunting Mitchell and Carson recommends eliciting individual WTP.

including CV research[6]. Mitchell and Carson's "household head" would, based on this model state the WTP that maximises the utility of the household. A feature of the unitary model is income pooling among household members, meaning that the source of income does not influence consumption decisions. The unitary model has come under fire in a large number of studies both in developing and developed countries (see e.g. Vermeulen (2002)), and is increasingly being replaced by what is known as collective models. Collective models allow for the fact that households often consist of several members who may have different preferences so that resource allocations are determined by some sort of cooperative (Pareto efficient) or non-cooperative bargaining among the members. In a typical CV survey a random household member[7] would be asked WTP for a (small) change in the provision of an environmental good in one of two main ways[8]:

(I) What is your maximum individual (or personal) WTP (on your own behalf)?; or

(II) What is your maximum WTP on behalf of your household (or your household's maximum WTP)?[9]

---

[6] This outcome results either from imposing a structure on the household decision-making problem so that the household utility function reduces to one (Samuelson), or through an altruistic (benevolent) head optimally allocating household resources (Becker).

[7] In some CV surveys the person identified as responsible for paying certain household expenses is taken to be the relevant decision maker, and answer on behalf on the household (e.g. like in Hensher et al. (2005)).

[8] The third way, not specifying the unit, as for example done in Delaney and O'Toole (2006), would in general not be advisable. Further, in English "you" which does not refer to an explicitly stated unit introduces an ambiguity in that there is no difference between plural and singular interpretation.

[9] We use the open ended WTP question format here, but what we write would naturally extend to other formats, such as dichotomous choice.

In CV studies, WTP from question I would normally be aggregated over individuals, and from II over households[10]. A commonly held view based on very little actual evidence, as pointed out by Strand (2007), is that the answer to II is higher than to I, but only if the respondent shows interpersonal preferences, such as altruism, towards the other household members. If there are no such preferences, respondents are seen to under represent household WTP, i.e. answers to I and II are (more) equal. In this case, aggregating WTP over households would grossly underestimate the total welfare change. But what does current theory and empirical results actually tell us to expect? To answer this question and clarify some of the confusion in the literature, it is useful to introduce some notation[11]. Let: hwtp = the maximum amount a household would be WTP for a change in the provision of the environmental good, so that all household members' utility levels are unchanged; $iwtp^i$ = household member i's response to WTP question I; and $hwtp^i$ = household member i's response to WTP question II.

In typical CV surveys, hwtp is unobservable, since what is observed is only a random household member i's representation[12], $hwtp^i$. Even if $hwtp^i$ is truthfully given, it may vary with which household member is asked (or misrepresent hwtp if given by a household head), and is therefore generally not equal to hwtp. For one-person households $iwtp^i$ should be equal to $hwtp^i$, which in this case equals hwtp by definition. The relationship between $iwtp^i$ and $hwtp^i$ on the one hand and hwtp on the other for multi-person households, is, however, not trivial. The most advanced attempt to investigate household decision making in this context is Strand (2007). He presents a collective model with no altruism assuming that the "true" hwtp is measured as the sum

[10] This may seem obvious, but we state it here since there is actually some confusion in the literature on this point.

[11] Inspired by Munro (2005).

[12] To the extent that the researcher can control who actually answers the survey.

of the adult household members' individual WTP (suppressing the summation index: $\sum iwtp_i$) i.e. each adult members' reply to WTP question I above[13]:

(1) $hwtp \equiv \sum_{i=1}^{m} iwtp^i \equiv \sum iwtp$, where the household has m (adult) members[14].

Strand's model assumes that the household allocates resources in efficient Nash bargains over a private and a household good (i.e. a good consumed commonly within the household). He then argues that WTP question I is most reasonably interpreted as member i's willingness to give up units of his privately consumed good for the increase in the public good. Question II, on the other hand, is interpreted as member i's trade-off between the household good and the public good. An implication of the model, is that in the bargaining solution individual i is always willing to give up more of the household good than of his private good, or:

(2) $hwtp^i > iwtp^i$

---

[13] In contrast to Strand, Quiggin (1998) and Bateman and Munro (2006) consider hwtp to be WTP obtained by consensus if adult household members are asked question II jointly. However, this intepretation makes implicit assumptions about the household bargaining structure allowing a consensus to be made in a stated preference survey setting. For example, it is unclear how consensus about hwtp could be achieved in a non-cooperative setting. Further, Bateman and Munro (2006: p3) state that "In one treatment a randomly selected individual is chosen from the couple and takes part in a face-to-face interview, *providing responses on behalf of the household"*. [our italics]. This is what B&M call "individual WTP", in our terminology $hwtp^i$, which they sum to an aggregate hwtp measure. In our setup, this would only make sense for $iwtp^i$ (equation (1)). However, this imprecision may be due the income pooling model assumed in their study, in which the distinction between WTP questions I and II becomes immaterial (see equation (5) below).

[14] Strand assumes in his basic model two adult household members only, but his results extend to more than two members, so we take the general case here.

The intuition for (2) is as follows. A small increase in common household income can be spent either on the private goods or the household good. In the bargaining solution, each member only receives a fraction to spend on the private good, and since in optimum the marginal consumption value of both types of goods must be equal, it follows that each member's marginal utility of the household good is lower than for the private good. (2) is the first theoretical prediction we will test in our data. Note that this result does not arise because of altruism, the generally held view mentioned above. Further, Strand (2007) derives the important result that individual i generally will misrepresent hwtp, in her answer to question II, i.e.

(3) $\text{hwtp}^{i} \neq \sum \text{iwtp}$

This result, which does not depend on bargaining strength between the members (or altruism), arises because the household members generally will have different marginal valuations of the public good (in terms of the household good). If member i's marginal valuation is higher (lower) than the other household member(s)', hwtp^{i} will instead be higher (lower) than $\sum$iwtp. Equality will only be achieved in the special case where marginal valuations are the same. However, importantly for practical CV research, Strand (2007: p541) argues that: "In a large random sample of households, such individual valuations should on average represent the respective households correctly, only provided answers are truthful". In other words the true, unobservable mean household WTP in a (large) sample ($\overline{\text{hwtp}}$), is equal to the observed mean response to question II ($\overline{\text{hwtp}^{i}}$):

(4) $\overline{\text{hwtp}} = \overline{\text{hwtp}^{i}} = \overline{\text{m}} \times \overline{\text{iwtp}^{i}}$

Given the definition of hwtp in (1), equality two in (4), where $\overline{m}$ is average (adult) household size and $\overline{iwtp^i}$ average response to question I in the sample from one random household member for each household[15], should also follow by approximation for large samples[16]. Equality should also hold if $iwtp^i$'s are added up with respondent-level household sizes instead of average household size. We cannot test (3) in our data, since we do not have observations of $iwtp^i$ from all members of the same household. However, through (4) we obtain our second prediction that can be tested. If (4) is supported, the implication is that asking WTP questions I or II should be immaterial to the welfare estimate, as long as the aggregation is done according to the chosen unit.

Strand (2007) then extends his model to a situation where the household members exhibit various forms of altruism toward each other. He shows that the bias in (3) is lower under pure[17] or paternalistic[18] altruism with respect to the public good than in the base case with no altruism. This generally contrasts with the well-known result derived within a unitary framework by Quiggin (1998). He finds that summing $iwtp^i$ to get a measure of hwtp leads to double counting as altruistic individuals when asked WTP question I generally fail to count the reduction in welfare for other household members[19]. However, Quiggin's result, as pointed out by Strand, coincides with a

[15] Note that only in a special case will $1/m\sum_{i=1}^{m} hwtp^i = hwtp$, i.e. the error in representation of household WTP for all members will in general not average out within the same household.

[16] The approximation of adjusting individual WTP with average household size is also used by Hasler et al (2008).

[17] Each household member attaches utility to the general utility level of the other

[18] Members care only about the other members' utility from consumption of the public good.

[19] Strand argues that one implication of his results, i.e. in (2), is that it leads to double counting to aggregate $hwtp^i$ values for each household member as a measure of hwtp, an implication he claims in Quiggin instead is driven by

special case where the individual exhibits non-paternalistic altruism, and is asked a version of question I where it is specified that only the individual will pay in reduced private consumption, and not the other household members – a distinction which matters in Strand's model framework[20]. Quiggin (1998) also shows that in the absence of altruism or when paternalistic altruism only arises with respect to the public good, the welfare measure will be invariant to the choice of response unit, i.e. (4) should hold by approximation[21]. Finally, within a fairly general framework that does not depend on the type of household allocation model, Munro (2005) shows that if and only if the household members pool income, household and individual WTP will be equal, i.e. the distinction between response units introduced with WTP questions I and II is unnecessary:

(5) $iwtp^i = hwtp^i = hwtp = hwtp^j = iwtp^j$, where members i and j are from the same household. The result follows directly from the properties of the indirect utility function once the utility functions depend on the sum of both individuals' income. Note that the result does not assume altruistic preferences. In this case, summing over individually stated WTP for the individuals in the sample would grossly exaggerate hwtp, in contrast with (4), as these are really representations of hwtp.

Empirically derived expectations

---

the assumption of altruism. However, Strand's argument is somewhat unclear since, in our view, it is $iwtp^i$ that is aggregated over household members in Quiggin and other studies, not $hwtp^i$.

[20] Quiggin only has a private and a public good in his model, and not the household good introduced by Strand.

[21] Quiggin's interpretation is not directly testable in our data, since he argues that $hwtp^i$ is likely to be a biased representation of hwtp, unless all members are asked jointly.

In addition to the theory above, economics combined with disciplines such as sociology and psychology offer empirical findings that can help us better understand the relationship between household and individual WTP. In a meta-analysis of CV studies of non-timber benefits Lindhjem (2007) finds mean individual WTP to be higher than household WTP, a counter-intuitive result from what we have discussed above, i.e.:

(6) $hwtp^i < iwtp^i$

Lindhjem (2007) suggests that this result may be due to the fact that reference to individual or household in the WTP question triggers different "mental accounts" (Thaler 1999; Li et al. 2005) or "psychological purses" (Webley 1995) from which the payment for the environmental good is drawn. These frequently observed psychological phenomena make money non-fungible in practice. Reference to the household may discipline the respondent's WTP as compared to the situation where he is thinking of his more generous private consumption budget. The issue becomes more complicated when considered within a household context (Winnett and Lewis 1995), but it is likely that such anomalies are even more common within households than for single individuals. A well-known finding in sociology is that men typically keep a higher share of their income for personal consumption than women do for the same income levels (e.g. Lundberg et al (1997)) so it is likely that if the mechanism underlying (6) is at work, it may be more commonly observed among men. Further, this literature finds that the degree of financial integration within a household, from complete income pooling to a separate finances, strongly influences household consumption decisions in general (Pahl 1995) and household vs. individual WTP in particular (Delaney and O'Toole (In Press)). Delaney and O'Toole (2004) find that a range of household and respondent characteristics explain how people perceive their agency when stating WTP for public

service broadcasting in Ireland, in response to a WTP question where unit is not specified. Being married, having children and being female are factors that suggest people respond as households rather than as individuals. Further, bargaining strength between members in a household, typically measured as relative income of partners, has been shown to influence household consumption decisions in various ways (Dosman and Adamowicz 2006). Strand (2007) shows that the difference in (2) above will be larger for members with low bargaining power.

The CV literature also suggests there are many respondent characteristics, often not directly derived from standard economic theory, that may explain variation in WTP. For example, preferences are sometimes shown to be different between women and men (Teal and Loomis 2000) and between people with or without children (Dupont 2004). Some of these variables can explain variation in WTP, but may importantly also, as pointed out by Delaney and O'Toole (2006), influence the self-perceived unit of response (even if it is explicitly stated in the WTP question) making the demand function endogenous. For example, higher WTP for women in a bid function may be because women tend to answer the WTP question more with the household in mind than as individuals. Including variables based more on empirical observations than well-established economic-theoretic relationships may help us better understand the relationship between household and individual WTP.

Testing procedure, survey design and data

Testing procedure

The theoretical predictions and empirically derived expectations discussed above were tested using a 2x2 split sample CV design (see Table 1 below). In sample A the respondent first got WTP question II ($hwtp^i$) and then prompted to instead think about

personal WTP she got WTP question I ($iwtp^i$). After the second WTP question, the respondent was automatically directed in the survey to a question offering 4-6 reasons for why $hwtp^i$ was higher, lower or the same as $iwtp^i$ (see next section, and Appendix for details). The design was the same in sample B, except the order of the WTP questions and the way the prompt was phrased were reversed (see next section).

Table 1 Split sample design testing differences in household and individual WTP

	Sample A	**Sample B**
1st WTP question	$hwtp_A^i$	$iwtp_B^i$
Prompt and 2nd WTP question	$iwtp_A^i$	$hwtp_B^i$
Post elicitation question about reasons for why $hwtp^i$ was higher, lower or the same as $iwtp^i$	Yes	Yes

The design allowed us to investigate the hypotheses discussed in the previous section (see Table 2 below for summary of these) comparing mean WTP within and between the two samples. Giving both questions I and II to the same respondents allows us to model the relationship between $hwtp^i$ and $iwtp^i$ using explanatory variables for household and respondent characteristics.

Table 2 Summary of testable hypotheses of mean individual and household WTP[1]

	Within samples (k=sample A, B)	**Split sample**	**Reference**
H1	$\overline{hwtp_k^i} > \overline{iwtp_k^i}$	$\overline{hwtp_A^i} > \overline{iwtp_B^i}$	Strand (2007). Eq. (2) Basic model
H2a[2] H2b[3]	$\overline{hwtp_k^i} = \overline{m_k} \times \overline{iwtp_k^i}$ $\overline{hwtp_k^i} = \frac{1}{n_k}\sum_{i=1}^{n_k} m_k^i \times iwtp_k^i$	$\overline{hwtp_A^i} = \overline{m_B} \times \overline{iwtp_B^i}$ $\overline{hwtp_A^i} = \frac{1}{n_B}\sum_{i=1}^{n_B} m_B^i \times iwtp_B^i$	Strand (2007). Eq (4) Response bias evens out in large samples
H3	$\overline{hwtp_k^i} = \overline{iwtp_k^i}$	$\overline{hwtp_A^i} = \overline{iwtp_B^i}$	Munro (2005). Eq. (5) Income pooling makes units equal
H4	$\overline{hwtp_k^i} < \overline{iwtp_k^i}$	$\overline{hwtp_A^i} < \overline{iwtp_B^i}$	Lindhjem (2007). Eq.(6) Explorative, based on mental accounting

Note: 1.Deriving the hypotheses it is reasonably assumed that the relationships discussed on the respondent level extend to means of samples. 2. It is somewhat artificial to test this hypothesis between samples for the second WTP question as people are expected to anchor their response to the first WTP question. 3. $iwtp^i$ is multiplied with respondent-level household size (m_k^i) and summed over the sample size (n_k).

CV survey design and environmental commodity

The data used for the experiment were collected from an Internet survey as part of a large multi-mode CV survey of forest protection in Norway. The government has currently protected ca. 1.4% of the productive forest area, which according to most biological assessments is too little to protect representative parts of forest habitats and endangered biodiversity. There are therefore plans to increase the level of protection, based on assessments of costs and benefits. The environmental benefits are primarily related to biodiversity (most of which are insects, fungi, mosses and plants) and ecosystem protection, though for some people the old-growth forest areas would increase recreation benefits. A professional polling firm collected the data in October 2007 from a pre-recruited nation-wide panel of respondents who had accepted to participate in surveys. The recruited respondents are informed by the polling company that they alone are supposed to answer the surveys, so there is a higher degree of control over who actually answers than can be expected from mail surveys[22]. The survey was designed following similar forest protection surveys well tested and tried in the Nordic context (see Lindhjem (2007)), and recent best-practice guidelines in the CV field (e.g. Bateman et al (2002), SEPA (2006)). The instrument went through thorough testing in focus groups and two small pilots (using both internet and personal interviews).

The survey first included questions about general use of government money for various ends to put the environmental good into a wider perspective, before focusing on the respondent's experience and use of forests in terms of recreational activities, and attitudes towards the perceived biological and aesthetical state of forests. Information

[22] However, it cannot be ruled out that some respondents consulted their partner or other household members when answering the survey on computers at home.

was then presented about number and types of species, and the interplay between forestry practices, protection and development of the ecosystem functions and biodiversity in forests. Six colour photos of endangered species and forest habitats were shown on the screen as well as pie and bar charts of number and percentage of species in all types of Norwegian habitats, including forests. The rather complex information was broken up with questions to activate the respondent and encourage response. Hard-to-avoid technical terms in the text (such as "biodiversity", "ecosystems", "endangered species", "nature reserve" etc) were explained in boxes that would pop up if respondents moved the cursor across underlined words. After this information, respondents were presented current forest protection policy (status quo) and future plans. The environmental commodity was specified as two forest protection plans of either an increase to 2.8% (doubling) or to 10% (possible long-term target), presented together[23]. The text was supplemented with digital, zoomable colour maps of current and future forest reserves, and a table giving information about the size of new reserves (in "mål"[24] og km^2), location of reserves (2/3 in Southern Norway[25]), and the improvements in the living conditions for main groups of species. The biological information was provided by a team of leading biologists in Norway, and checked by foresters to ensure a balanced presentation of the status quo and future plans.

Household and individual WTP elicitation and follow-up probes

[23] The full survey varied the second level of protection for each respondent between 4.5% and 10% for the purpose of conducting an external scope test, reported elsewhere.

[24] Traditional area of measurement in Norway equal to 1/10 hectare.

[25] Since the exact location of future reserves is not yet decided the existing reserves in Southern Norway were increased in size on the maps to the correct relative size (2.8% and 10% of productive forest area, respectively), to give the respondent an idea of space requirements.

The basis for the comparison of hwtp^{i} and iwtp^{i} was the 2.8% protection plan. After the information about the two plans, the respondent was given the following text (the bold was varied between Samples A and B):

"We ask you now to consider how much the two alternative plans are worth for **your household/you**. Think through carefully how much the 2.8% plan is worth compared to the current situation, before you give your final answer to the next question. Try to consider what would be a realistic annual amount given **your/your household's** budget. **You/Your household** must choose whether to spend the amount on the forest conservation plan, or on other things. What is the most **your household/you** almost certainly **is/are** willing to pay in an additional annual tax earmarked to a public fund for increased forest conservation from today's level of 1.4% to 2.8% of the productive forest area? Choose the highest amount, if anything, **your household/you** almost certainly will pay".

People could then indicate their maximum WTP in a payment card (PC) in the form of a drop-down menu with a non-linear scale containing 24 amounts (ranging from 0 first to NOK 15000[26]), including "don't know" (at the end). The amounts where chosen on the basis of previous CV studies (e.g. Lindhjem (2007)). PC was chosen as response format over dichotomous choice, to avoid yea saying and anchoring (at the expense of theoretical incentive compatibility) (Boyle (2003)). PC also lends itself nicely to the drop-down menu format very familiar to internet-users. The payment vehicle (an earmarked tax to a forest protection fund) was chosen because it is response unit neutral, for example compared to an income tax (a potential problem in e.g. Hasler et al

[26] There was also an option to choose "more than 15000", in which case a box would pop up where the exact amount could be specified.

2008), is realistic and reduces people's scepticism that the money would not be spent on forest protection. The typical budget reminder, included in most modern CV studies, referred either to "your (personal)" or "your household's budget". Following the first WTP question, the respondent would get the following prompt, before getting the second WTP question with response unit changed: "We now ask you if it matters for your willingness to pay if you state it **for yourself/on behalf of your household** or **on behalf of your household/for yourself**."

After the second WTP question, respondents were automatically taken to a follow-up question asking whether a number of stated reasons were important, not important or not relevant for their response to the two questions. The suggested reasons allow the respondent to express her considerations regarding her (and her partner's) preferences (for herself, family), budget (individual, household, common), and role in the household (e.g. usually paying household expenses). The respondent could also state openly other reasons that may have been important. The details of these are given in the Appendix, and the descriptive statistics of peoples' responses given in the next section. The rest of the CV survey followed standard procedure, probing into why people answered zero or positive, checking their understanding and perceived realism of the scenario and WTP questions. The final part collected socio-economic background information, which was merged with existing panel information about the respondents and their households.

Results and analysis

Mean individual vs household WTP between and within samples

Before estimating mean WTP the dataset was cleaned. Around 10% of respondents in both samples (and for both WTP questions) chose the "don't know" option in the drop-

down menu for the PC, while between 20-25% chose zero. There were no significant differences in these answers across samples. Since our main aim here is to investigate the relationship between household and individual WTP and people's stated reasons, all zeros and "don't know" responses were therefore removed[27]. Further, respondents from one-person households were taken out. This procedure reduced the samples from around 400 to 240 observations, each. A comparison of mean values of sample characteristics indicated no immediate reason for applying weighting procedures or using covariates in the estimation of mean WTP (see Table 7 in the next section). Since the stated WTP amounts had a skewed distribution with a long right tail, a log-transformation of WTP was applied. Mean WTP for the interval PC data for the two samples and WTP questions were estimated following standard procedures given in Cameron and Huppert (1989) (see Table 3 below)[28] .

---

[27] A very small number of respondents answered either "0" or "don't know" on one of the WTP questions and a positive number on the other, which suggests that such responses generally were made for reasons unrelated to the wording of the question as household or individual WTP. Further, it may me regarded as inconsistent to answer a positive WTP on the individual question, and "0" on the household question (while the opposite would be possible since the respondent may think other household members have positive WTP). Twice as many men as women answered "0" (while "don't know" was equally distributed), though the share was the same between samples.

[28] We compared a normal and lognormal model with a simple non-parametric survival function using the lower bound of the PC intervals. The lognormal model showed a better fit. Mean WTP from this model is given by $E(WTP)=\exp(a+\sigma^2/2)$, where a and σ are the estimated parameters from the lognormal model.

Table 3 Mean annual individual and household WTP for the two samples in Euros

	Sample A		Sample B	
WTP Questions	**Mean**	**95% CI**	**Mean**	**95% CI**
1st	$\overline{\text{hwtp}_A^i} = 172$	(141, 203)	$\overline{\text{iwtp}_B^i} = 154$	(121, 188)
2nd	$\overline{\text{iwtp}_A^i} = 147$	(121, 173)	$\overline{\text{hwtp}_B^i} = 237$	(182, 292)
N	239		234	

Note: Estimated using interval regression in STATA 9.2. Confidence intervals were calculated using 1000 bootstrap draws with replacement, following Efron (1997). 1 Euro = 8.07 Norwegian Kroner at time of study.

The response to household WTP (Euro 172) is higher than to the individual WTP (Euro 152) between the samples for the first question as expected from theory, confirming the common view in the CV literature. However, the difference seems not to be significant at the 5% level. We ran a likelihood-ratio test to check statistical significance, see (7):

$$(7)\quad q = -2\left[\log L_{\text{PooledAB}} - (\log L_A + \log L_B)\right] \sim \chi^2(\text{d. f.})$$

where $\log L_A$ and $\text{Log}L_B$ refer to the log likelihood values of from the estimated models for individual samples, and $\log L_{\text{PooledAB}}$ is the likelihood value for a pooled model. Running the pooled model without a sample dummy, yields a test static $(\hat{q})$ of 6.96, which allow us to reject that both parameters are equal at the 3% level. However, running the same model with a sample dummy yields $\hat{q}$=2.12, which means we cannot reject that the standard errors are the same at 10% level (i.e. the samples can therefore be pooled). The dummy is significant at 2.8% level, indicating confirmation of the one-sided hypothesis that $\overline{\text{hwtp}_A^i} > \overline{\text{iwtp}_B^i}$. However, an extended bootstrap (10000 draws with replacement) from each of the sample distributions combined with a simple non-parametric test of means indicates ca 80% $\overline{\text{hwtp}_A^i} > \overline{\text{iwtp}_B^i}$ and 20% $\overline{\text{hwtp}_A^i} \leq \overline{\text{iwtp}_B^i}$. This means that we can reject the hypothesis that household WTP is higher than

individual WTP between the samples for the first WTP question[29]. The confidence intervals estimated around the means in Table 3, also indicate that equality cannot be rejected at the 5% level. For the second WTP question where response units are reversed, respondents in Sample A generally reduce their bids (mean Euro 147) while respondents in sample B increase their bids (mean Euro 237), as expected. However, this difference is not symmetric, as can be seen from Table 4 below:

Table 4 Percent of respondents who answered higher, the same or lower on household WTP compared to individual WTP question

	Sample A	**Sample B**
Higher ($hwtp^i > iwtp^i$)	32.6	52.9
Same ($hwtp^i = iwtp^i$)	59.4	44.4
Lower ($hwtp^i < iwtp^i$)	7.9	2.5
Total	100%	100%

Around 53% of sample B increase their bid from $iwtp^i$ to $hwtp^i$, while only 33% reduce their bids from $hwtp^i$ to $iwtp^i$ in sample A. The reason for this stickiness downwards is not immediately clear. It is possible that some people in Sample A interpreted the $hwtp^i$ question as an $iwtp^i$ question (despite the unit being explicitly stated), and therefore saw no reason to reduce their bid in the second question (similar to what was found in Delaney and O'Toole In press)[30]. Another explanation, drawing parallels to psychology and the embedding debate in the CV literature where e.g. parts and the whole of a good is valued, is that people who value the whole good first and than a part of it, tend to keep (or only slightly reduce) their bid faced with new information (Veisten et al. 2004).

[29] This arises because sample B's distribution has a lower mean, but a higher standard error than sample A's distribution, effects partly outweigh each other in the formula for mean WTP (see previous footnote).

[30] It was not possible to control whether respondents would take the time and trouble to go back a page in the Internet survey to change their response to the first WTP question, when confronted with the second question. We included a control sample that only got one $hwtp^i$ question. Mean WTP from this sample was almost identical with that of Sample A, indicating that this practice was not prevalent in the survey

Pairwise t-tests on the difference of bootstrapped mean WTP values between WTP questions I and II within each sample were conducted, strongly confirming $\overline{\text{hwtp}^i_k} > \overline{\text{iwtp}^i_k}$ for both samples at the 1% level[31].

To check hypotheses H2a&b, we also scaled individually stated WTPs with the number of adult household members[32] from each respondent's household and the average household size in the samples (see Table 5). For the former case, high and low limits from the PC data were multiplied by the household size, and interval regressions rerun. For the latter case, previously estimated CI was scaled by the constant average household size for each sample.

Table 5 Individual WTP in Euros scaled up by measures of household size

Mean adjustment factor	**Sample A**	**Sample B**
Average household size in sample[1]	340	382
Adult members of each respondent's household	345	387
N[2]	224	218

Note: 1. Mean household sizes sample A: 2.29, Sample B: 2.41. 2.The sample sizes are lower than in Table 3 since the respondent database had some missing values about household size for a few respondents.

Interestingly, the two methods to adjust individual WTP with household size yield very similar results. Compared to the household WTP for sample A in Table 3, adjusting individually stated WTP estimates to represent household WTP, yield significant overvaluation. $\overline{\text{hwtp}^i_A}$ is about half of the individual WTP adjusted by household size

[31] 95% CI around the mean for this test using bootstrap was estimated at Euros (172, 174) for $\overline{\text{hwtp}^i_A}$, (147, 149) for $\overline{\text{iwtp}^i_A}$, (154, 156) for $\overline{\text{iwtp}^i_B}$ and (236, 240) for $\overline{\text{hwtp}^i_B}$.

[32] Adult members include people above the age of 15. Our dataset did not allow us to easily filter out adults between the age of 15 and 18 living with their parents, which would maybe have been the optimal solution. Since we have screened out all one-person households, the average household size will naturally be somewhat higher here than in the general population.

from sample B, contrary to the expectation in equation (4). Also for the second WTP question where the majority end up answering $hwtp^i > iwtp^i$, there is no equality of mean $hwtp^i$ and mean adjusted $iwtp^i$ within samples. Finally, we can summarize our empirical results in Table 6 below.

Table 6 Summary of empirical results by hypothesis (k=sample A,B)

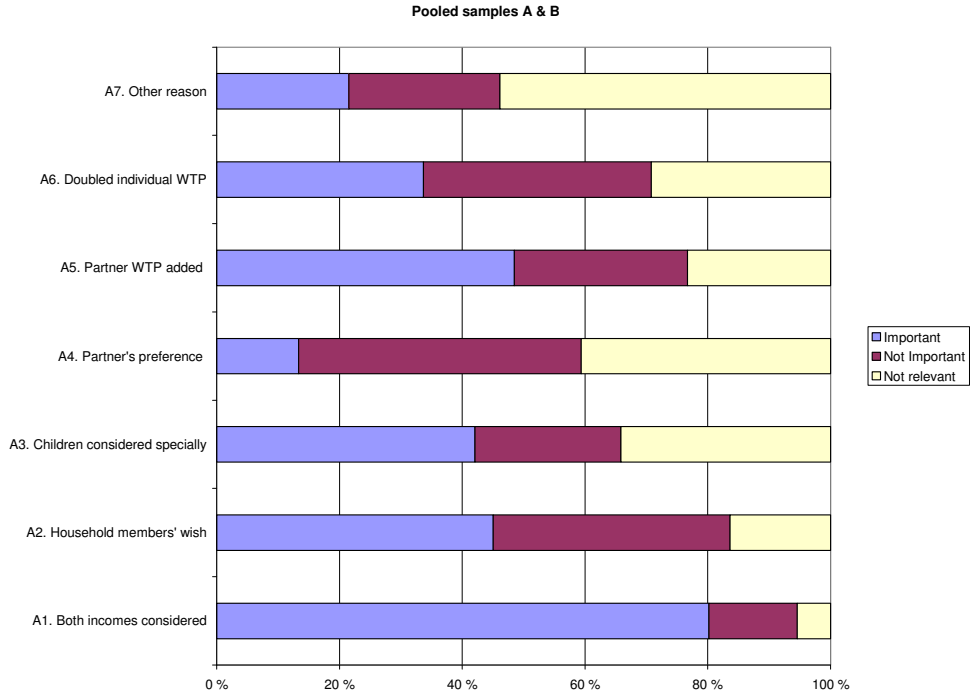
	Within samples	**Test result**	**Split sample**	**Test result**
H1	$\overline{hwtp_k^i} > \overline{iwtp_k^i}$	**Supported A&B**	$\overline{hwtp_A^i} > \overline{iwtp_B^i}$	Rejected
H2a	$\overline{hwtp_k^i} = \overline{m_k} \times \overline{iwtp_k^i}$	Rejected A&B	$\overline{hwtp_A^i} = \overline{m_B} \times \overline{iwtp_B^i}$	Rejected
H2b	$\overline{hwtp_k^i} = \frac{1}{n_k}\sum_{i=1}^{n_k} m_k^i \times iwtp_k^i$	Rejected A&B	$\overline{hwtp_A^i} = \frac{1}{n_B}\sum_{i=1}^{n_B} m_B^i \times iwtp_B^i$	Rejected
H3	$\overline{hwtp_k^i} = \overline{iwtp_k^i}$	Rejected A&B	$\overline{hwtp_A^i} = \overline{iwtp_B^i}$	**Supported**
H4	$\overline{hwtp_k^i} < \overline{iwtp_k^i}$	Rejected A&B	$\overline{hwtp_A^i} < \overline{iwtp_B^i}$	Rejected

Rejection of H1 and H2 between samples follows from the support to hypothesis H3. However, household WTP is significantly higher than individual WTP within each sample (i.e. H1 supported), where rejection of H3 and H4 (but not H2) within samples logically follows.

Explaining the relationship between individual and household WTP

We now turn to trying to explain the observed relationship between individual and household WTP. Some explanations are given by the respondents themselves, when stating in the follow-up question for each proposed reasons whether it was important, not important or not relevant to their choice. Figures 1 and 2 sum up the results for the pooled samples, for $hwtp^i > iwtp^i$ and $hwtp^i = iwtp^i$, respectively. Full versions of statements respondents considered are given in the Appendix.

Figure 1 Percentage of respondents rating given reasons for $hwtp^i > iwtp^i$ (n=202)



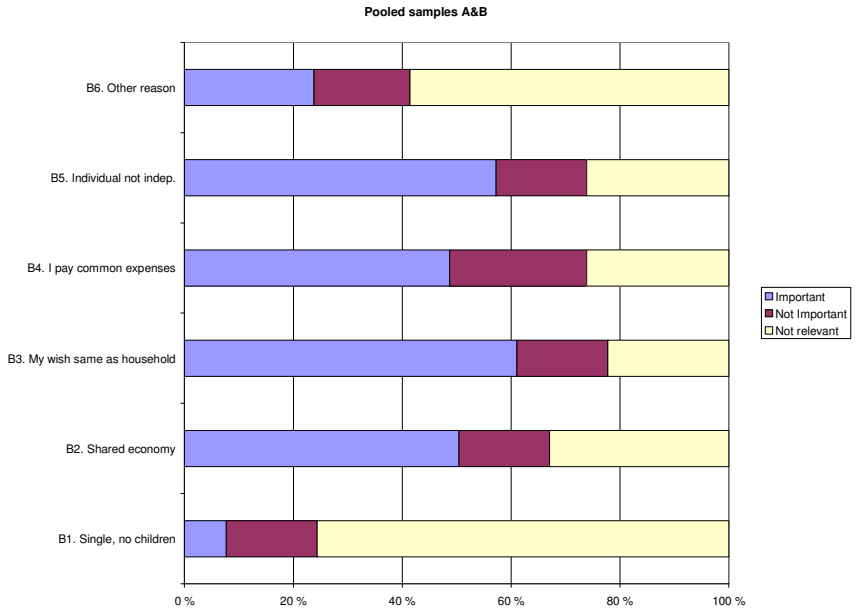
The by far most important reason why people state higher household WTP, is that they have a larger budget at their disposal and therefore can pay more (80% for reason A1 in Figure 1). This means that individuals do not seem to consider the income of adult co-members as part of their own budget constraint, i.e. income is not pooled. The second most important reason (49%) is that the estimated WTP of a partner is added (A5), which may be consistent with both a separate and shared economy in practice. The third reason (A2) is that respondents think about the household also when answering individually (i.e. no generous individual "mental accounts" distort the expected relationship between $hwtp^i$ and $iwtp^i$). The fourth most important reason is that respondents consider children especially when answering the household WTP question (A3). This is an indication of some kind of altruistic preferences, which is the traditional view of why household WTP may be higher (but as discussed not a necessary condition in Strand's (2007) model). Only 33% state that they have just doubled the individual amount since they are from a two-adult household (A6). Finally, it is interesting to note

that very few respondents have answered higher household WTP because their partner have a stronger preference (higher marginal valuation) for forest conservation than themselves (A4). There were no substantial differences in answers to this follow-up question between samples A and B. However, reasons A1 and A5 seemed to be more important for sample B, and A6 for sample A (all ca 10 percentage points). Some of the open answers (A7) pointed to the number of household members (with or without reference to use of forests) as important, that individual WTP was calculated as a percentage, that assumptions had been made about other members' WTP, and generally that a larger budget is available (supporting A1).

A similar battery of reasons where offered for $hwtp^i$=$iwtp^i$ respondents (see Figure 2). The most important reason is that the respondent's wish concurs with what the household would collectively have decided (61%, reason B1 in Figure 2). This is an indication of a unitary household model (as judged by the respondent), as is the reason almost as many states as important: taking household members into account even when answering as an individual (57%, B5). About 50% indicate as important that they have a shared economy with their partner (B2), 48% that they are responsible for paying household expenses and therefore that it does not matter for their WTP which response unit they are asked (B4). Reason B3 was indicated as more important for sample A than B (by 12 percentage points), while other reasons had no big difference across samples. In their open statements, some respondents had noted that the amount they decided on the first question was "an appropriate amount" or "enough", i.e. that they saw no reason to increase or decrease from this amount. This is a reason related to the embedding argument discussed above; once the respondents have decided on an amount, they will not change in light of new information. Hence, in our dataset there may be a few such

responses, where "stickiness" ($hwtp^i$=$iwtp^i$) is not due to reasons related to household decision-making.

Figure 2 Percentage of respondents rating given reasons for $hwtp^i$ =$iwtp^i$ (n=234)[33]



Some respondents also indicated that they made the decisions in their household, i.e. similar to the Mitchell and Carson's "household head", and therefore $hwtp^i$=$iwtp^i$. Only 7.9% in sample A and 2.5% in sample B (total of 25 respondents) chose $hwtp^i$<$iwtp^i$. Around 50% of these indicated as important that their personal budget was higher (reason C1 in Appendix), i.e. indicating "mental accounting" (or possibly separate finances). One respondent mentioned separate finances as important, another that it was easier to answer the individual question than assuming WTP for the other members.

[33] One-person households without children were taken out. Still, 7% had indicated that being single and having no children (B1) was important. This may be due to the fact that the CV data generally are more updated than the Internet panel information. Further, some of the responses may be due to misunderstanding or "random answers", in any case a low number giving us some degree of confidence in people's responses.

We also posed the question of whether household and respondent characteristics, defined in Table 7 below, can explain the relationship between individual and household WTP. We do not rule out a priori that a range of respondent variables (e.g. sex, age, education, use, attitudes, etc) often included in CV bid functions (see e.g. Banzhaf et al (2007), may also be important to explain the relationship between household and individual WTP. These variables may be considered explorative. Of household variables we included number of household members, presence of children, altruistic attitudes for WTP>0 (latter two crude proxies for altruism), if children (>15 years) answer the survey, type of residence, whether grocery purchases are jointly planned, marital status (latter two crude proxies for economic integration), and the respondent's share of household income (common proxy for bargaining strength). We chose a simple approach with a binary dependent variable of 1 if $hwtp^i > iwtp^i$ and 0 if $hwtp^i = iwtp^i$ estimated using a standard probit model, similar to the approach in Delaney and O'Toole (2004). The few respondents answering $hwtp^i < iwtp^i$ were excluded for simplicity. Models using the WTP ratio ($hwtp^i/iwtp^i$) or difference ($hwtp^i - iwtp^i$) as dependent variables instead of the binary variable, were specified and tested, but gave generally lower explanatory power. Being unfamiliar with the task, it is likely that respondents had a clearer idea about the direction than the exact magnitude of the difference between $hwtp^i$ and $iwtp^i$. Using a regression model we can control for the different characteristics that are underlying people's responses and hidden in Figures 1 and 2 above.

Table 7 Explanatory variables and sample means (st. dev.) samples A and B

Variables	Definition	A	B
Respondent characteristics:			
Sex*	Dummy: 1 if male, 0 if woman	.504	.512
		(.032)	(.032)
Age*	Continuous: >15 years	41.2	41.7
		(.91)	(.96)
Incomeind	Individual income 2006, Norwegian Kroner 25000 intervals	316912	312948
		(11903)	(13597)
Eduhigh*	Dummy: 1 if > 4 years university education; 0 if mid-education	.100	0.119
		(.019)	(.021)
Edulow*	Dummy: 1 if only primary education; 0 if middle education	.075	.085
		(.017)	(.018)
Owner	Dummy: 1 if forest owner; 0 if not	.306	.235
		(.029)	(.027)
Member	Dummy: 1 if member of nature org.; 0 if otherwise	.025	.038
		(.010)	(.012)
Use	Dummy: 1 if forest used for recreation last 12 months; 0 if not	.924	.935
		(.017)	(.016)
Highuse	Interaction variable: 1 if >15 times in forest last month and	.063	.136
	"Use"=1; 0 otherwise	(.015)	(.022)
Nouse	Dummy: 1 if sure not to use proposed forest reserves, 0 if otherwise	.172	.179
		(.024)	(.025)
Attax*	Dummy: 1 if agree that high taxes ensure public goods; 0 if otherw.	.138	0.16
		(0.02)	(0.02)
Altruism	Dummy: 1 if respondent indicated as reason for WTP>0 that other	.277	.238
	people can enjoy old growth forests; 0 if otherwise	(.029)	(.027)
Household characteristics:			
Relinc	Individual income as share of household income	.570	.545
		(.015)	(.017)
Childdum*	Dummy: 1 if children <15 years of age in household; 0 if otherwise	.231	.227
		(.027)	(.028)
Childresp*	Dummy: 1 if son/daughter answered questionnaire; 0 otherwise	.077	.080
		(.017)	(.018)
Married*	Dummy: 1 if married; 0 if previously married/single	.596	.607
		(.032)	(.032)
Cohabit*	Dummy: 1 if cohabitants; 0 if previously married/single	.236	.209
		(.027)	(.027)
Grocery*	Dummy: 1if divided responsibility, grocery purchase; 0 if otherw.	.454	.446
		(.032)	(.033)
House*	Dummy: 1 if detached house; 0 if otherwise	.592	.638
		(.032)	(.032)
Hhldmem*	Number of adults and children (1-4, 5 or more)	2.98	3.03
		(.071)	(.071)
N**		239	234

Note: * Variable information taken from Internet panel of respondents. Other variables are from the CV survey. ** Some averages based on reduced sample. No weighting was conducted between samples.

Results for the separate and pooled samples are displayed in Table 8 below. The models show reasonable fit to the data, but coefficient estimates should be interpreted with caution. For sample A, older people have significantly higher probability to state equality, indicating perhaps both for long relationships and for the older generation, the difference between the individual and the household gets increasingly blurred.

Table 8 Probit models on $hwtp^i > iwtp^i$ (Y=1) or $hwtp^i = iwtp^i$ (Y=0) for separate and pooled samples

Independent variables	Sample A		Sample B		Pooled sample (A+B)	
	Coefficient	**Z-score**	**Coefficient**	**Z-score**	**Coefficient**	**Z-score**
Dummy for sample (WTP question order)					.482***	3.69
Respondent variables:						
Sex	-.158	-0.73	.387*	1.73	.126	0.85
Age	-.025**	-2.34	.004	0.44	-.006	-0.94
Incomeind	0.000	1.48	0.000**	1.96	0.000***	2.59
Eduhigh	.166	0.55	.265	0.89	.193	0.96
Edulow	.412	0.96	.670	1.36	.534*	1.76
Owner	-.370	-1.62	.127	0.55	-.013	-0.09
Member	-.439	-0.51	.249	0.44	.214	0.51
Use	.521	1.44	-.040	-0.10	.259	0.99
Highuse	-.404	-0.97	-.120	-0.43	-.145	-0.67
Nouse	-.073	-0.29	-.622**	-2.38	-.405**	-2.34
Altruism	.163	0.71	.354*	1.66	.198	1.32
Attax	-.095	-0.34	-.503*	-1.90	-.375**	-2.00
Household variables:						
Relinc	-1.362***	-2.16	-2.209***	-3.68	-1.653***	-4.01
Childdum	.453	1.27	-.251	-0.81	-.064	-0.29
Childresp	1.150*	1.66	.253	0.39	.611	1.38
Married	.842	1.59	-.200	-0.48	.202	0.65
Cohab	.945*	1.86	-.151	-0.36	.348	1.13
Grocery	-.173	-0.84	.005	0.03	-.048	-0.36
House	.054	0.26	-.242	-1.04	-.103	-0.71
Hhldmem	-.374**	-2.50	-.022	-0.18	-.104	-1.14
Constant	.948	1.11	.944	1.11	.273	0.48
Log Likelihood	-120.19		-129.82		-260.84	
Pseudo R^2	0.1329		0.1419		0.1228	
N	214		219		433	

Note: ***, **, * Indicates significance at 1%, 5%, and 10%, respectively

The other respondent variables are not significant for sample A. In sample B, men have a significantly higher probability than women of answering $hwtp^i > iwtp^i$, i.e. women may see income pooling as more natural than men. Individual income level has no effect (not significant in sample A)[34]. Interestingly, people who are certain not to visit the future forest reserves (variable "Nouse") have lower probability of separating between household and individual WTP (not significant for sample B). It is likely that non-users are more likely to misinterpret the $iwtp^i$ question as a $hwtp^i$ question than user

[34] It may have been more appropriate to categorise the income variable in a few groups, than running the models with continuous amounts, though results may not have been very different.

with familiarity of the good, in which case the WTP difference will be smaller (see e.g. Delaney and O'Toole In press). Another explanation, if indeed respondents have clearly understood the response unit, is that a "moral dump" or donation has been made, which is less sensitive to unit of response (and, as is often found, to other key dimension of the survey instrument, such as scope of the good)[35]. People who favour a tax to pay for public goods ("Attax") also display smaller WTP differences.

For the household characteristics, only relative income is significantly negative through all three models. A higher share of household income reduces the probability of stating $hwtp^i > iwtp^i$, as expected. This result is consistent with Strand's (2007) model of bargaining strength, but can simply also be interpreted as an indication of separate finances as people's WTP is strongly correlated with their personal budget constraint. In sample A, cohabitants and married have a higher probability of answering $hwtp^i = iwtp^i$, compared to the base case of single people (with children) and previously married, suggesting financial integration. Shared responsibility for grocery purchases suggests the same (though not significant). Increasing number of household members (i.e. more than 2) reduces the probability of answering $hwtp^i > iwtp^i$, through all models (only significant for A)[36]. This suggests that once children are involved, the household is more tightly integrated, resulting in smaller differences between household and individual WTP. This result runs contrary to the common argument that altruism drives a wedge between individual and household WTP. If the alleged effect of altruism is present, it may be outweighed by the higher degree of income pooling that emerges

[35] This is more often found in simple CV surveys than in the more detailed and well-tested surveys that have folloeed best-practice guidelines. We therefore think the first explanation is more probable for our survey.

[36] Mere presence of children in the household (i.e. "Childdum") did, however, not have a consistent and significant effect on response probability.

once relationships mature. A few children (>15 years) answered the survey, and these generally had a higher probability of answering $hwtp^i > iwtp^i$ ("Childresp") as expected since they have lower income than their parents. A final point to note is that the dummy on WTP question order (i.e. sample) is significantly positive indicating a higher probability of $hwtp^i > iwtp^i$ in sample B. The results of the models give us some degree of confidence in the validity of the data, and supplement the insights provided by respondents in their stated reasons for differences in household and individual WTP.

Discussion and conclusions

The practical implications for aggregate welfare estimates of the choice of response unit for WTP – household or individual – in contingent valuation surveys has been largely ignored in the literature to date. In this paper we demonstrate that the empirical consequence may be substantial noise or bias in welfare estimates. In our CV survey of forest protection in Norway, following standard protocols for design and testing, we find that people state a higher mean WTP when asked on behalf of the household than as individuals in a split sample test, but this difference is not significant at the 5 % level. Aggregating WTP over individuals in this case more than doubles the total welfare change from forest protection in Norway compared to mean stated household WTP. This result runs contrary to Strand's (2007) collective household decision model, which predicts that response unit biases would even out over large samples, i.e. response unit is immaterial as long as mean WTP is aggregated consistently over the unit of choice. Since we only include respondents with positive WTP and multi-person households in our experiment, the distortion in welfare estimates would be somewhat lower for a full sample.

When people are prompted in the second WTP question to answer for the other response unit, an average of 43% decide to state higher household WTP than individual WTP, while 52% state the same WTP. More people state higher household WTP if they have been asked individual WTP first, i.e. people tend to more easily increase their bids than reduce them. Mean household WTP within the same samples is found to be significantly higher than individual WTP on the 1% level. 80% of respondents state as an important reason for this result that they have a larger budget at their disposal when asked household WTP. There are few indications that altruism, though imperfectly measured in this study, may be important in explaining that household WTP is higher than individual WTP within samples – the commonly held view in the literature (as described by Strand (2007)). Instead, degree of financial integration and relevant budget constraints seem to be more important.

Our study is a first attempt to investigate the empirical differences between individual and household WTP for one type of environmental good in a particular CV setting – where both types of questions could meaningfully be asked. The degree to which our results can be generalised to other CV studies, or stated preference research more generally, types of goods, survey modes etc is uncertain. For other environmental goods of a more quasi-private nature, such as hunting permits, individual WTP questions may be the appropriate choice. However, more empirical research is undoubtedly necessary within stated preference valuation, to advance the theory of intra-household resource allocation and to test it empirically. Recent research interviewing partners and households together in choice experiment settings (such as in Bateman and Munro (2006), Beharry and Scarpa (2006)) are important contributions. However, since interviewing partners or households together never really will be a practical option due to excessive costs, empirical work should inform stated preference design, where

random individuals typically are asked, with the aim to reduce response unit distortion in welfare estimates as much as possible. An important point made by Delaney and O'Toole (2006, In press) is that people's self perceived agency – i.e. their interpretation of the unit of the WTP question notwithstanding explicit reference to "household" or "individual" – may vary depending on household and respondent characteristics. We think the risk of such misunderstandings to some extent may be alleviated by carefully designing the survey to be consistent with the chosen response unit. Not only wording of the actual WTP question, but the type of good (e.g. extent of non-use values), the payment vehicle (e.g. household tax vs income tax), budget reminders and scenario descriptions, and even the survey mode[37], may give conflicting cues as to the intended agency of the respondent. Our results indicate that people may need more information to state their WTP reliably for the household or as individuals. Some sort of advance notice before the WTP questions (which is recommended e.g. for scope tests – see e.g. Bateman et al (2004)), "cheap talk" and definition of "household" may be useful in clarifying the intended response unit. These approaches will have to be carefully tested not to introduce other, unintended biases, a common experience in the history of CV. However, our results suggest that response unit distortions may be sufficiently problematic to need fixing. In the meantime, asking household WTP for environmental goods of random individuals is the most conservative approach and should be followed, even though it may lead to underestimation of welfare change in some situations.

[37] People may for example view their agency differently depending on whether they are asked during an intercept at a forest site or in a shopping mall compared to filling in a paper questionnaire on their own, family kicthen table.

Appendix

Table A: Given reasons in survey for why respondents chose higher, the same or lower household or individual WTP

How important were the following reasons for you stating a higher/same/lower amount on behalf of your household than/as for yourself? **Answers: Cross "Important, Not Important, or Not Relevant"**		
Higher	**Same**	**Lower**
A1. I took both incomes into account when I was asked on behalf of the household	B1. I am single and have no children, so there is no difference	C1. I thought about my individual budget and can pay more than if I have to take my household into consideration
A2. I take the household members' wishes regarding increased forest conservation into account even if I consider willingness to pay for myself alone[1]	B2. My partner and I have a shared economy, so it does not matter if I am asked personally or on behalf of my household	C2. My partner is against more forest conservation, so I adjusted for that
A3. I especially consider the children when asked on behalf of the household	B3. What the household collectively would have decided concur with my wish	C3. I am normally not the one paying for our household expenses, so I chose a lower amount on behalf of my household
A4. My partner is more interested in forest conservation than I am, so I adjusted for that	B4. I am normally the one paying our household expenses, so in practice there is no difference if I am asked personally or behalf of my household	C4. We have a tight budget for household expenses, but my personal budget is more generous
A5. I added what I think my partner would be willing to pay	B5. I take my household members into account even if I consider willingness to pay for myself alone	
A6. I doubled my individual amount since we are two adults in the household		
Other reasons that were important? Specify:_____		

Note: 1. There was some ambiguity in the interpretation of this reason. The intended meaning is that the respondent does not focus on a more generous individual budget "mental account" for $iwtp^i$, but thinks about the whole household and therefore goes up from $iwtp^i$ to $hwtp^i$.

Acknowledgements

The authors would like to thank Jon Strand, University of Oslo and IMF, Olvar Bergland, Norwegian University of Life Sciences, Anna Bartczak, Warsaw University, and Mark Morrison, Laura O. Taylor and other participants at the Environmental and Resource Economics Researcher Workshop, 12-13. November 2007, Charles Sturt

University, Australia for constructive comments. Funding from the Norwegian Research Council (grant 173230/I10) is gratefully acknowledged.

References

Arrow K J, Solow R, Leamer E, Portney P, Radner R and Schuman H (1993) Report of the NOAA Panel on Contingent Valuation. Federal Register 58: 4601-4614.

Banzhaf H S, Burtraw D, Evans D and Krupnick A (2006) Valuation of natural resource improvements in the Adirondacks. Land Economics 82(3): 445-464.

Bateman I, Cole M, Cooper P, Georgiou S, Hadley D and Poe G L (2004) On visible choice sets and scope sensitivity. Journal of Environmental Economics and Management 47: 71-93.

Bateman I and Munro A (2003) Non-cooperative decision-making and measures of household surplus

Bateman I and Munro A (2006) Household versus individual valuation: what's the difference? Kyoto, Japan

Bateman I J, Carson R T, Day B, Hanemann W M, Hanley N, Hett T, Jones-Lee M, Loomes G, Mourato S, Ozdemiroglu E, Pearce D W, Sugden R and Swanson T (2002) Economic Valuation with Stated Preference Techniques: A Manual. Edward Elgar Publishing, Cheltenham, 480pp.

Becker G S (1973) A theory of marriage: Part I. Journal of Political Economy 81: 813-846.

Becker G S (1981) A Treatise on the Family (Enlarged edition). Harvard University Press.

Beharry N and Scarpa R (2006) Joint vs separate decisions by couples in choice experiments: The case of coastal water quality in Tobago. Kyoto, Japan

Bergstrom T C (1994) A survey of theories of the family, University of Michigan

Bergstrom T C (2003), Benefit Cost Analysis and the Entanglements of Love. PROCEEDINGS OF SESSION II: HOUSEHOLD DECISION MAKING.

Boyle K J (2003) Contingent valuation in practice. In: P. A. Champ, K. J. Boyle and T. C. Brown, eds, A primer on nonmarket valuation. Kluwer Academic Publishers

Cameron T A and Huppert D D (1989) Ols Versus Ml Estimation Of Non-Market Resource Values With Payment Card Interval Data. Journal of Environmental Economics and Management 17(3): 230-246.

Carson R T, Flores N E and Meade N F (2001) Contingent valuation: Controversies and evidence. Environmental and Resource Economics 19(2): 173-210.

Delaney L and O'Toole F (2004) Irish Public Service Broadcasting: A Contingent Valuation Analysis. The Economic and Social Review 35: 321-350.

Delaney L and O'Toole F (2006) Willingness to pay: individual or household? Journal of Cultural Economics 30: 305-309.

Delaney L and O'Toole F (In press) Individual, household and gender preferences for social transfers. Journal of Economic Psychology.

Dosman D and Adamowicz W (2006) Combining stated and revealed preference data to construct an empirical examination of intrahousehold bargaining. Review of Economics of the Household 4: 15-34.

Dupont D P (2004) Do children matter? An examination of gender differences in environmental valuation. Ecological Economics 49(3): 273-286.

Efron B (1997) Bootstrap methods: Another look at the jacknife. Annals of Statistics 7: 1-26.

Hasler B, Jacobsen J B, Lundhede T H, Martinsen L and Thorsen B J (2008) Household versus Individual Willlingness to Pay: Evidence from a Choice Experiment on Environmental Valuation

Hensher D, Shore N and Train K (2005) Households' Willingness to Pay for Water Service Attributes. Environmental & Resource economics 32: 509-531.

Li H, Berrens R P, Bohara A K, Jenkins-Smith H C, Silva C L and Weimer D L (2005) Testing for budget constraint effects in a national advisory referendum survey on the Kyoto protocol. J. Agric. Resour. Econ. 30(2): 350-366.

Lindhjem H (2007) 20 Years of stated preference valuation of non-timber benefits from Fennoscandian forests: A meta-analysis. Journal of Forest Economics 12: 251-277.

Lundberg S J, Pollak R A and Wales T J (1997) Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit. The Journal of Human Resources 32(3).

Mitchell R C and Carson R T (1989) Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington DC.

Munro A (2005) Household willingness to pay equals individual willingness to pay if and only if the household income pools. Economics Letters 88(2): 227-230.

Pahl J (1995) His money, her money: Recent research on financial organisation in marriage. Journal of Economic Psychology 16: 361-376.

Pollak R A (2003) Gary Becker's contributions to family and household economics. Review of Economics of the Household 1: 111-141.

Quiggin J (1998) Individual and household willingness to pay for public goods. American Journal of Agricultural Economics 80: 58-63.

Samuelson P A (1956) Social Indifference Curves. The Quarterly Journal of Economics 70(1): 1-22.

SEPA (2006) An instrument for assessing the quality of environmental valuation studies. Swedish Environmental Protection Agency,

Smith V K and Van Houtven G (1998) Non-Market Valuation and the Household, RFF

Smith V K and Van Houtven G (2004) Recovering Hicksian consumer surplus within a collective model: Hausman's method for the household. Environmental and Resource Economics 28(2): 153-167.

Strand J (2005) Individual and household values of mortality reductions with intrahousehold bargaining. Journal of Risk and Uncertainty 31(2): 217-236.

Strand J (2007) Public-good valuation and intrafamily allocation. Environmental and Resource Economics 38(4): 527-543.

Teal G A and Loomis J (2000) Effects of gender and parental status on the economic valuation of increasing wetlands, reducing wildlife contamination and increasing salmon populations. Society and Natural Resources 13(1): 1-14.

Thaler R H (1999) Mental accounting matters. Journal of Behavioral Decision Making 12(3): 183-206.

Veisten K, Hoen H F and Strand J (2004) Sequencing and the adding-up property in contingent valuation of endangered species: Are contingent non-use values economic values? Environmental and Resource Economics 29(4): 419-433.

Vermeulen F (2002) Collective household models: principles and main results. Journal of Economic Surveys 16(4): 533-64.

Webley P (1995) Accounts of accounts: En route to an economic psychology of personal finance. Journal of Economic Psychology 16: 469-475.

Winnett A and Lewis A (1995) Household accounts, mental accounts, and savings behaviour: Some old economics rediscovered? Journal of Economic Psychology 16: 431-448.