University of Warsaw
Faculty of Economic Sciences

Working Papers

No. 3/2008

Mikołaj Czajkowski
Nick Hanley

How to 'Sell' an Environmental Good: Using Labels to Investigate Scope Effects

Warsaw 2008

University of Warsaw
Faculty of Economic Sciences

Working Paper No. 03/2008

How to 'Sell' an Environmental Good: Using Labels to Investigate Scope Effects

Mikołaj Czajkowski
Warsaw Ecological Economics Center
Faculty of Economic Sciences, University of Warsaw
miq@wne.uw.edu.pl

Nick Hanley
Department of Economics
University of Stirling

Abstract

Insufficient sensitivity to scope remains one of the pivots of criticism addressed at validity of stated preference methods. Many studies demonstrate failure of a scope test of some sort, while many others show that WTP responses are sensitive to the scope of environmental change. Despite some existing explanations and reasons for insensitivity to scope (embedding, warm glow, uncertainty over supply of a public good, awareness of all options) there seems to exist no clear conclusion on how to deal with it. The paper provides an alternative explanation for insufficient sensitivity to scope, based on redefinition of value drivers of environmental goods. In the proposed framework respondents' WTP need not depend only on physical characteristics of a valued good but may also partly be a function of a 'label' under which the environmental good is 'sold'. To investigate this problem and empirically test the hypothesis and its implications, a CE study in a biodiversity valuation setting is designed and conducted. The conceptual framework and empirical evidence provide an alternative explanation for problems with the insufficient sensitivity to scope observed in many studies. Finally, we set out some implications of the nature of labels as value drivers for the design of future valuation studies.

Keywords:
scope test, embedding, warm glow, label effect, CE, CVM, biodiversity

JEL:
Q51, Q57, Q23

Introduction

Stated preference methods remain the only source of estimates for both use and non-use values, and hence can provide valuable inputs to cost-benefit analyses of environmental change. Despite the initial distrust many researchers held for such methods, they have eventually entered the "mainstream" of economic science (Carson and Hanemann 2005), and are routinely used as part of the policy analysis process. However, certain important issues remain to be resolved in terms of enhancing the validity of welfare estimates. These issues have arisen in the context of testing whether the results of valuation methods are in line with the predictions of economic theory.

There are at least two such tests. Firstly, it is expected that as the price increases the demand for an environmental good should decrease. For instance, in dichotomous choice CVM designs, this means that the ratio of the respondents who answer 'yes' to a bid question should decrease with increases of the bid amount. A second test consists of examining the prediction that respondents should be willing to pay more as the amount or quality of environmental good to be provided increases. Failure to pass this *scope test* has been traditionally one of the pivots of criticism of stated preference methods (Goodstein 1995). The purpose of this paper is to investigate sources of failure to pass the scope test. An alternative explanation is proposed, based on redefinition of the value drivers of environmental goods. Empirically, we show that respondents' WTP might depend not only on physical characteristics of a good being valued, but partly also on the 'label' under which the good is being 'sold'. This

provides an alternative explanation for problems with the scope test observed in many studies: failing to control for the effects of a label leads to a mis-interpretation of scope differences.

In the next section, we briefly review the literature on scope tests, and then offer an alternative explanation in terms of labels. Section 3 describes the design of an empirical study to investigate this idea, whilst section 4 contains results. The final section concludes.

1. Explaining Scope Effects, or their absence

2.1 Existing Literature

As noted above, the scope test consists in testing whether respondents are willing to pay statistically more for a larger amount of an environmental good, in terms of increased quantity or quality. Even though the origins of this test can be tracked earlier, most of the debate was based on the critique put forward by Kahneman and Knetsch (1992).

There exist two versions of scope test – internal and external. In an internal version the same respondents are asked to state their WTP for different levels of environmental good improvements. In an external version – two different levels are valued by different respondents using split sample.[1] There seem to be less problems with internal scope tests (Brookshire et al. 1976; Smith and Osborne 1996); some studies argue, however, that this might be due to the urge of the respondents to maintain 'internal integrity' of their answers (Heberlein et al. 2005). On the other hand, according to Adamowicz et al. (1999) an internal

[1] Alternatively, Bateman and Brouwer (2006) test scope sensitivity comparing the WTP declared for a respondent and his entire household.

scope test allows for a comparison of pair wise WTP estimates of each respondent in the sample, and thus controls for the heterogeneity of the respondents. There seems to be no clear consensus in the literature whether insufficient sensitivity to scope is a regularly observed phenomenon, or happens occasionally, or on how its causes might be resolved and eliminated. Some of the best known examples of failing the scope test are given by Kahneman (1986), Kahneman and Knetsch (1992), Diamond and Hausman (1994), Hausman (1993), Cummings (1989), Schwartz (1997), Svedsäter (2000) and Shiell and Gold (2002). On the other hand there are meta-analyses indicating that the scope test is usually passed successfully (Walsh et al. 1992; Smith and Osborne 1996; Carson 1997; Brouwer et al. 1999; Rosenberger et al. 1999; Poe et al. 2000).

Scope test might be failed due to one of the following reasons: (1) insufficient power of the test, taking into consideration the difference in the level of provision of a public good (Arrow and Leamer 1994); (2) errors resulting from invalid construction of hypothetical scenario which masks the true, underlying sensitivity to scope (Carson and Mitchell 1995); (3) embedding (Kahneman and Knetsch 1992); and (4) the warm glow effect (Becker 1974). Other explanations offered in the literature include uncertainty over supply of a public good (Powe and Bateman 2004) and lack of respondent's full awareness of all the possible alternatives (Bateman et al. 2004). Carson *et al.* (2000) analyze possible errors in construction of valuation scenarios that might lead to scope insensitivity. Some of the reasons include unclearly defined goods or changes in the level of their provision (Carson and Mitchell 1993). In this case a respondent would be valuing a different good from that intended by the researcher. This might also be one of the manifestations of embedding.

Another possible reason for insufficient sensitivity to scope is the *warm glow effect*. As Knetsch and Sinden (1984) noted, respondents stating their WTP for public or environmental goods may in fact 'purchase moral satisfaction'. The idea, inspired by the Olson's (1965) concept of 'inpure altruism', was later renamed by Becker (1974) as a warm glow effect. The concept results from a notion that a respondent in fact derives additional utility from contributing to a public good, due to social assent, prestige or moral satisfaction.[2] The concept was later used by Andreoni (1989; 1990) to demonstrate theoretically why progressive taxes might increase social and charitable spending and why government spending on these does not crowd out private contributions. Accepting warm glow as an explanation for insufficient sensitivity to scope would mean that a welfare change resulting from implementation of a particular scenario would not be fully and directly associated with changes to the good in question.[3] This would imply a need to take into account motives for wanted a public good to be supplied when estimating values (McConnell 1983; 1997; Bergstrom and Reiling 1998; Johansson-Stenman 1998).

Even though warm glow or purchasing of moral satisfaction might contribute to insensitivity to scope (Cooper et al. 2004) there is a major problem with accepting these explanations for all observed problems with scope test. Despite the fact that warm glow effects are relatively well described in the literature devoted to charitable and social organizations, its extrapolation to the valuation of public goods would necessary mean that consumers derive higher utility with an increase in the taxes they pay or bids they accept (Chilton and Hutchinson 1999). This would mean that there should be a positive relationship between declared WTP values and

[2] According to some authors, in case of public goods the existence of warm glow effect means that respondents are in fact expressing their ethical, rather than economic preferences (Kahneman and Ritov 1993; Ritov and Kahneman 1997).

[3] This might be crucial in some applications such as cost-benefit analysis. If the respondents' WTP partly reflect their 'warm glow', the welfare estimates are problematic to be used directly.

utility level derived solely from 'moral satisfaction' reasons. However, virtually all studies (to our knowledge) implementing warm glow as a component of utility function do this by introducing a non-zero constant representing warm glow in utility functions if a non-zero WTP is declared. This constant is obviously independent of WTP or bid level.

A conclusion one can draw from this brief overview of the literature on scope effects is that more research seems to be required in order to broaden understanding of the reasons for scope insensitivity. The next section offers an alternative explanation which does not depend on warm glow motivations.

2.2 A new approach to thinking about scope effects

We put forward a hypothesis that the elicited value of an environmental good depends not necessarily only on the physical characteristics of the good in question, but also on the 'label' under which it is 'sold'. The label can be an important attribute in itself, in accordance with the decomposition framework put forward by Lancaster (1966). However, a label is different from all the other attributes because it is independent from all the physical (quantifiable) characteristics of the good, and depends instead on the respondent's perception regarding the brand. This notion is in line with framing dependence, as suggested by (Kahneman and Tversky 2000). We demonstrate that this approach allows for an alternative explanation of scope test problems and is devoided of the weaknesses of the warm glow effect.

There is a vast marketing literature devoted to the importance of labels and brands, and how their associated images influence choices. Consumers are known to have preferences over

labels in addition to the physical characteristics of goods. The influence of labels on consumers' utility and choices is also supported by neurological studies (Roe and Haab 2007). A stark example is provided by McClure et al. (2004), who replicated 'The Pepsi Challenge' using Functional Magnetic Resonance Imaging (fMRI) of subjects' brains to investigate preferences and information processing in choices between two well known soft drinks presented with and without brand names. The respondents given two samples of the same soft drink supplied without brand names (labels) were essentially indifferent. However, when one of the two identical samples was labelled as Coke or Pepsi, subjects systematically preferred Coke to the unlabelled alternative (which they were told could be Coke or Pepsi), despite both drinks being chemically identical (both Coke or both Pepsi). These results were also reflected in different neural responses of the subjects, who seemed to have neurologically processed labels (brands) differently. This provides a stark manifestation of labels influencing preferences. The neural evaluative process, and subsequent choice of the product, may thus be altered by the presence of a label.

We argue that a value of an environmental good (and probably, in general, a public good) elicited using stated preference methods depends partly on its physical characteristics and partly on the label under which it is presented to the respondents. To formalize the approach let indirect utility function v be a function of a price vector P, income y, and a vector of attributes of an environmental (public) good Q (for simplicity assume one public good).

$$v(P, y, Q) \tag{1}$$

The attributes of the good consist of a label q_L and a set of remaining physical (quantifiable) attributes Q_{-1}. Assume q_L to be a binary variable representing inclusion of label in the description of the good, and the label to be desirable (thus $\frac{\partial v}{\partial q_L} > 0$). In addition, since labels are independent from physical characteristics of the good we assume q_L to be additively separable from Q_{-1}.

$$v\left(P, y, q_L, Q_{-1}\right) \tag{2}$$

This formulation allows us to derive compensating and equivalent surplus measures, where superscript 0 denotes the initial, and 1 the post-change level of the good:

$$\begin{aligned} v\left(P^0, q_L^0, Q_{-1}^0, y^0\right) &= v\left(P^0, q_L^1, Q_{-1}^1, y^0 - CV\right) \\ v\left(P^0, q_L^1, Q_{-1}^1, y^0\right) &= v\left(P^0, q_L^0, Q_{-1}^0, y^0 + EV\right) \end{aligned} \tag{3}$$

In a standard random utility setting (McFadden 1974), assuming a vector of socio-demographic and choice specific explanatory variables Z, respondent's j indirect utility becomes:

$$v_j\left(P_j, q_{L,j}, Q_{-1,j}, y_j\right) = V_j\left(Z_j, q_{L,j}, Q_{-1,j}, y_j\right) + \varepsilon_j \tag{4}$$

where ε_j is unobservable, individual specific error term. Note that the label component of utility is also individual-specific, since individuals might differ in their perception of the significance or meaning of the label used. Assuming in turn additivity of the deterministic and stochastic parts of indirect utility function and an IID property of error terms, we may derive respondent's j willingness to pay[4] from:

[4] Assuming an improvement of an environmental good this will be their compensating surplus.

$$V_j\left(Z_j, q_L^1, Q_{-1}^1, y_j - WTP_j\right) + \varepsilon_j^1 = V_j\left(Z_j, q_L^0, Q_{-1}^0, y_j\right) + \varepsilon_j^0 \quad (5)$$

In the case of a linear WTP function the result might be then simplified to show that estimated mean WTP $\left(\overline{WTP}\right)$ would be a function of two additive components – the set physical characteristics of the good Q_{-1}, and the label which was used for describing the good q_L:

$$\overline{WTP(Q)} = g\left(Q_{-1}\right) + h\left(q_L\right) \quad (6)$$

The above formulation of mean WTP has some important implications. When two different levels of change in an environmental good are offered to respondents in a CVM exercise, and they are described using the same label, it is obvious that the estimated values of WTP would differ only with respect to its physical attributes Q_{-1} (the real changes, for example the percentage of habitat safeguarded from development, or the area of woodland to be planted), while the label component q_L of the mean WTP would remain constant. As a result, if the share of a value of the label in total WTP $\left(\overline{WTP}\right)$ was "sufficiently high", then the observed WTP estimates for two levels of environmental good would not be "sufficiently different". Only after controlling for the label effect on WTP would we expect the estimated WTPs in two, alternative change scenarios to pass the scope test. This is an idea which is amenable to empirical testing.

2. Design of the Empirical Study

We employed a labelled Choice Experiment (Blamey et al. 2000) to test if welfare measures of implementing an environmental policy incorporates a value of a label, which is not associated with any physical characteristics of the good. We then demonstrate how the value

of such a label can contribute to insufficient sensitivity to scope when alternative quantities of increase in the good are valued.

The context of the empirical application is the protection of forest biodiversity. The case study site was the Białowieża Forest in Poland, which is considered to be one of the most important remaining temperate natural lowland forests in Europe. Initial steps in identifying attributes and a label involved focus groups and verbal protocols, with subjects representing both the general population of Poland and local communities neighbouring the Białowieża Forest. Currently National Park designation applies to 16% of the Białowieża Forest area, and there is an ongoing debate about whether Park designation should be extended to the whole area of the forest. We found that respondents have preferences for extending the national park *per se*, irrespectively of what the extension would mean in terms of on-the-ground protection and management. Thus, national park designation was chosen as the label for providing environmental change, since it was not associated with any characteristics of the good to be valued (forest biodiversity), and yet appears to be widely recognized and desired by respondents.[5] It should be clearly stated that qualitative analysis conducted via focus groups and verbal protocols did not show any biodiversity attributes to be associated with the selected label. The label was found not to increase probability or quality of the provision itself or influence the scenario in any other observable way. It was thus safe to conclude that emotional associations of the label were not tied to any of the specific attributes of the good.[6]

In order to describe the likely changes in biological diversity of the forest, the most important elements of biodiversity were identified with cooperation from biologists and ecologists. This

[5] It's worth noting that many of the national parks in Poland have different management regimes and protection goals; thus 'extending national park' is not associated with any specific set of actions or characteristics.
[6] Such untied associations are referred to as 'freestanding emotions' (Rossiter and Percy 1997).

list of candidate attributes was then pre-tested on the general public using focus groups, resulting in the final selection of 4 attributes, at 3-4 levels. The selected attributes represented potential changes in the biodiversity as viewed by ecologists, as well as being comprehensible to respondents. The first attribute – *natural ecological processes* – represented natural dynamics of the Białowieża Forest. This illustrated natural changes of the forest's flora – processes of succession and regression, fluctuation, degeneration and regeneration, as well as seasonal changes. According to our specialists, and as explained in the questionnaire, improvements in this attribute could be achieved by passive protection of a given percentage of the total area of the Białowieża Forest. Three possible levels of this attribute were: *status quo* – 16% of the area protected, *partial improvement* – 30% protected, and *substantial improvement* – 60% of the area to be passively protected.

Rare species of fauna and flora represented the second attribute. It was underlined in the description that this attribute represents not only known, but also yet unknown species. Examples of both flagship and lesser-known species were provided, together with information concerning the likelihood of yet-unknown species occurring in the forest, and their dependence on protection. A short general explanation of the importance of different species to ecosystem was provided. The possible levels of this attribute were: *status quo* – a decline threatening total extinction, *partial improvement* – maintaining current populations, and *substantial improvement* – maintaining and expanding current populations. This attribute is rather similar to that included by Christie et al (2006) in their choice experiment on biodiversity protection.

Ecosystem components was the attribute characterizing the existence of biotopes and ecological niches, such as dead wood, natural ponds, streams, clearings etc. It was explained

in the questionnaire that improvements in this attribute may be achieved by active protection of these components. This attribute could be important for respondents both for the existence of the components alone, as well as a proxy for better well-being of species inhabiting the forest. The possible levels of this attribute were: *status quo* – the lack of some components and decrease in the quality of the existing ones, *minor improvement* – regeneration of deteriorated components across 10% of the forest area, *partial improvement* – regeneration and protection across 30% of the forest, and *substantial improvement* – regeneration and protection across 60% of the forest area.

The last attribute was monetary, representing the cost of an additional compulsory tax to be paid for the following 10 years. Each respondent was presented with four choice sets, each consisting of three alternatives. The first alternative always represented a status quo option and there was no variance in attribute levels (including cost). The two other alternatives were designed using orthogonal fractional factorial design. The first of these was always labelled 'extension of the national park' while the second was labelled 'other forms of protection'. Focus group studies and pretesting clearly showed that there was no embedding regarding 'extension of national park' since the respondents did not think that the extension itself would bring about any other changes than the ones described by the attributes; it was none-the-less strongly highlighted in the questionnaire that both alternatives would essentially provide the same changes to the environment. The design consisted of 32 choice sets blocked into 8 questionnaire versions. An example of the choice card is given in Figure 1. For a more detailed description of the study we refer the reader to Hanley and Czajkowski (forthcoming).

3. Results

The face-to-face CE surveys were conducted in June 2007 on a nationwide representative quota sample of adult Poles by a professional surveying company. There were a total of 400 surveys collected, resulting in 1600 choice observations.

The statistical analysis of CE was conducted using NLOGIT 4.0. A number of different model specifications were tried, including Multinomial Logit, Error Components, Nested Logit, Heteroscedastic Extreme Value, Random Parameters Logit and Multinomial Probit – each in many possible functional forms. In virtually all models we found statistical significance of "PARK", which represented choice of the labelled alternative – providing protection in a particular way. It is worth noting, that this variable represented value independent from changes in all the other attributes used. It may be interpreted as a premium consumers get when the protection plan is implemented through the extension of a national park – and thus providing an environmental good with a 'label'.

To select the best model the unnested models were compared using the Vuong test (Vuong 1989) and Clarke's distribution-free test (Clarke 2003; 2007). Where necessary, tests were corrected for different numbers of estimated coefficients using Schwarz's (1978) Bayesian information criterion. The tests showed that Nested Logit Model with normalization of the scale parameter at the top level of the tree outperformed other models. Detailed results of the final model are given in Table 1. The explanatory variables are dummies representing different possible improvements in the levels of the attributes, thus allowing for nonlinear marginal utilities (the status quo being the reference level). Because there was no statistical

difference between *partial* and *substantial* improvement in the *Rare species of fauna and flora* attribute, in most of the models the two levels are jointly represented as an *improvement*. PARK is a dummy representing the alternative specific constant for the labelled alternative 'extension of the national park' and *Cost* is the monetary variable measured in PLN[2007]. As noted above, the significance of the PARK variable shows the presence of a clear labelling effect.

Table 1. Nestel Logit Model estimates of the CE

Variable	Coefficient	Standard Error	*p-value*
Natural Ecological Processes (1-level improvement)	0.28213***	0.09479154	0.0029
Natural Ecological Processes (2-level improvement)	0.39521***	0.12019442	0.0010
Rare Species (improvement)	0.19484**	0.08553986	0.0227
Ecosystem Components (1-level improvement)	0.26939**	0.10561570	0.0108
Ecosystem Components (2-level improvement)	0.30377***	0.10600908	0.0042
Ecosystem Components (3-level improvement)	0.34398***	0.11899022	0.0038
PARK (alternative specific constant)[7]	0.34743***	0.06257316	0.0000
Cost	-0.02126***	0.00299983	0.0000
Inclusive value[8]	0.60569***	0.09979989	0.0000

***, **, * represent significance at 1%, 5%, 10% level respectively

Number of observations	1213
Log likelihood function	1220.485
Chi squared	204.0790
Degrees of freedom	9
$\Pr(\chi^2 > \text{critical value}) =$	0.0000000

[7] ASC *PARK* was representing particular way of implementing the change – providing the changes in the form of national park extension.

[8] The inclusive value for the non-restricted branch of the tree. The value between 0 and 1 is well within constraint for common component of random terms (Hensher and Greene 2002).

Implementing the approach suggested in Louviere et al. (2006), WTP values for each level of the attributes were calculated, with reference to the *status quo* level of each attribute. The results, given in euro[9], are summarized in Table 2. Standard errors were calculated using the Delta method; confidence intervals were estimated using parametric bootstrapping (Krinsky and Robb 1986).

Table 2. Implicit prices of the CE attribute levels (EURO)

Attribute	Implicit price	90% C. I.	Standard error	*p-value*
Natural Ecological Processes (1-level improvement)	3.69***	3,42 - 5,86	1.0842	0.0007
Natural Ecological Processes (2-level improvement)	5.16***	2,25 - 5,42	1.2913	0.0001
Rare Species (improvement)	2.55**	3,46 - 6,73	1.0215	0.0127
Ecosystem Components (1-level improvement)	3.52***	1,17 - 3,83	1.2586	0.0052
Ecosystem Components (2-level improvement)	3.97***	1,91 - 5,19	1.2335	0.0013
Ecosystem Components (3-level improvement)	4.49***	2,37 - 5,57	1.3055	0.0006
PARK (alternative specific constant)	4.54***	2,81 - 6,19	0.9468	0.0000

In order to demonstrate the influence of the label on scope sensitivity two policy scenarios were considered, both of which are currently being considered for actual implementation. In the first scenario (LO) only minimal improvements to all the attributes were included, while in the second scenario (HI) the attributes would be provided at the highest levels used in the CE design. The components of the two policies are summarized in Table 3.

[9] The values in euro were calculated using the following exchange rate: 1 euro ≈ 3.6 PLN.

Table 3. Components of the policy scenarios

Attributes	Policy scenario ' *LO* '	Policy scenario ' *HI* '
Natural Ecological Processes	1-level improvement	2-level improvement
Rare Species	improvement	improvement
Ecosystem Components	1-level improvement	3-level improvement

The resulting mean welfare estimates – $\overline{WTP_{LO}}$ and $\overline{WTP_{HI}}$ respectively – were calculated using the approach provided by Hanemann (1982). In order to compare the influence of the label on welfare estimates of the two policies in terms of scope sensitivity, the welfare estimates of the two policies were calculated including the label (Table 4a) and excluding it (Table 4b), through the device of either including or omitting the parameter on PARK in the compensating surplus calculations. Finding a scope effect would mean rejecting the null hypothesis that the mean WTP for the "low" scenario was equal to the mean WTP for the "high" scenario, since the scenarios differ in terms of the quantity of biodiversity conservation on offer. The means, standard errors and confidence intervals are based on parametric bootstrapping (Krinsky and Robb 1986).

Table 4a. Welfare estimates for policy scenarios with the label[10] (EURO)

Policy	**Welfare estimate**	**90% C. I.**	**Standard error**	***p-value***
LO	14.29	12.20 - 16.39	1.6347	.0000
HI	16.74	14.47 - 19.02	1.7764	.0000

[10] Providing environmental changes in the form of extending the national park.

Table 4b. Welfare estimates for policy scenarios without the label[11]

(EURO)

Policy	Welfare estimate	90% C. I.	Standard error	*p-value*
LO	9.75	8.04 - 11.46	1.3367	.0000
HI	12.20	10.37 - 14.03	1.4309	.0000

We first analyze the difference in means between the two policy scenarios differing in attribute levels applying Park et al.'s (1991) method of non-overlapping confidence intervals. Based on the bootstrapped empirical distributions of mean welfare estimates of policies *LO* and *HI* we conclude that the *p-value* that would allow to conclude that the welfare estimates are different (significance level assuring confidence intervals not to overlap) are 0.236 when the label is included, and 0.188 when the label effect is excluded. According to Poe et at. (1994; 2005) however, using confidence intervals for testing difference in means is inappropriate. Instead they propose a convolutions method, which consists in calculating all the possible differences between elements of a vector bootstrapped from the distribution ofthe higher mean and a vector bootstrapped from the distribution of the lower mean. As a result, the ratio of number of outcomes that are less than zero to the number of all the possible outcomes (equal to product of lengths of the two vectors) gives an exact *p-value* of the hypothesis that the mean of vector 1 is higher than the mean of vector 2. This is referred to as complete combinatorial convolutions method. Following this approach we have estimated *p-values* for the hypotheses that $\overline{WTP_{HI}}$ is higher than $\overline{WTP_{LO}}$ for the estimates including the

[11] Providing environmental changes without extending the national park.

label and excluding it.[12] The *p-value* for the null hypothesis that the welfare estimate of the policy scenario '*HI*' is equal to that for the policy scenario '*LO*' was estimated to be 0.1493 for the case of including the label, and 0.0996 with the label excluded. In other words, we can reject the null hypothesis of insensitivity to scope at the 10% level once labelling effect is controlled for.

This comparison demonstrates a potential influence that including a label in welfare estimates may have on sensitivity to scope of environmental policies. The two analyzed policies did not result in significantly different welfare estimates if the label effect was present in both of the welfare estimates (as the *p-value* is close to 0.15). However, excluding the label effect substantially increased the significance level of the difference. In our case excluding the label allows the researcher to accept the null hypothesis of difference in welfare estimates of the two policies at the 10% significance level: in other words, a statistically significant scope effect is only found when the label effect is controlled for. In the following sections we analyze the implications of these findings.

4. Discussion

Our empirical study demonstrated that WTP for an increased protection level of an environmental resource may carry an additional component, resulting from presenting the scenario using a label which is recognized by the respondents as desirable. This may be the case even if the label does not associate with any identifiable attributes left out from the CE

[12] 1e5-element vectors of welfare estimates were bootstrapped for each of the four policy scenarios. Estimating all the possible differences of pairs of vectors resulted in two 1e10-element vectors. Number of negative elements of each vector was calculated and divided by total number of elements to derive *p-values*.

design. Such a 'label effect' may be of a more general character, and may apply to a range of environmental and public goods that can be associated with any sorts of labels. The value of the label can indeed constitute an important share of total value (in our case roughly 30%). Thus our empirical study confirms that the value of label may be a considerable component of the total WTP. This validates the relationship made explicit in formula (6), that WTP might depend not only on the physical attributes of the good, but may also consist of a constant component associated with the value of a label.

In our case the label represented providing additional protection in the particular way – extending the national park. The preferences of respondents for national parks, even if no particular protection regime or attributes are associated with them, seem to be confirmed by the results of Bartczak et al. (forthcoming) and Jacobsen and Thorsen (forthcoming). These authors also observe premiums for the existence of national parks *per se*, i.e. irrespective of the actual protection policy implemented there.

Our results clearly show that controlling for the value of a label in welfare estimates substantially increases the sensitivity of these estimates to changes in the physical characteristics of environmental goods (scope). This conclusion seems to be general, as long as (1) a label constitutes an important share of total economic value and (2) covariances between the parameter of a label in the indirect utility function with parameters of other attributes are non-negative. If (2) does not hold then excluding the label from welfare estimates would cause both the mean and standard error to decrease, which could result in uncertain impacts on the level of significance of the difference between welfare estimates of the two policies.

The theoretical framework of our hypothesis allows for explaining the insufficient sensitivity to scope observed in some studies. Since virtually all CV studies employ a specific scenario for valuation one may expect that at least in some of these studies some form of label may be an implicit but un-identified component of WTP. This would inevitably cause welfare estimates to be less sensitive to changes in scope of the valued good. It is not clear whether this would matter more for an attribute-based valuation like choice experiments compared with a whole-valued based approach such as contingent valuation. Important practical questions that arise are what might constitute a label for an environmental policy or good, and how to identify and devise the appropriate label for valuation scenarios. We believe that there are many possible labels depending on the particular good and on respondents' preferences. Identifying appropriate labels requires qualitative analysis – pretesting, focus groups or verbal protocols, in a manner no different from selecting all the other attributes for the study.

The value of a label is directly connected to the good in question or the scenario of its provision. It is thus worth discussing if the value of the label should or should not be included in welfare estimates. The answer to this question seems to depend on the research goal – if the study is aiming at estimating the total change of consumers' welfare due to a change in provision of an environmental good, it is safe to conclude that the value of the label should be included in the total economic value estimates of the scenario. This represents a notion of consumer sovereignty (Carson et al. 2000) – if 'purchasing' a good with a particular label generates additional utility it should be accounted for in welfare estimates. On the other hand, if the aim of the study is to estimate implicit prices of the changes in environmental qualities (marginal or discrete) – it should probably be excluded and only the value resulting from the

changes of physical attributes should be included. As our case study demonstrates, this is possible using the CE method.[13]

Our formulation of total economic value consisting of, among the others, a label, is partly in line with the works of Rolston (1988), Gunn (1980) and Rescher (1980) who propose the concept of 'intrinsic' value of the species, that would be irrespective of its actual representatives (real animals). This, according to them, explains the apparent paradox, that an animal of an endangered species would be more valuable than an animal of the same species except that not endangered. This view was criticized by Russow (1981) who claimed that the definition of a species is somewhat flexible and so species *per se* cannot have values, because defining more species would simply increase the total value. Assigning part of the value to labels recognized by the respondents deals with these problems and explains apparent paradoxes. This is because there may be only as many 'premiums' for species as there are recognizable labels. The value of a label of the same species may also be substantially different between two countries, depending on a role the species plays in national identity for instance (Jacobsen et al. 2008). An interesting example that can be interpreted as an importance of the label is provided by Jacobsen et al. (2008) who observe different welfare estimates of conservation programmes when describing them using quantitative descriptions of results, than when describing them using the names of the species to be protected. Even if the respondents were not familiar with the names of the species to be protected they seemed

[13] This discussion is somewhat similar to including / excluding ASC in welfare estimates of the scenarios (Adamowicz et al. 1998). Among many studies Rolfe et al. (2000), Bennett et al. (2001), Horne et al. (2005), Horne (2006) and Birol et al. (2006) included ASC in welfare estimates. On the contrary – Xu et al. (2003), Lehtonen et al. (2003), Biénabe and Hearne (2006) and Nielsen et al. (2007) include only implicit prices of physical attributes. Finally Adamowicz et al. (1998), Garber-Yonts et al. (2004), Watson et al. (2004), Mogas et al. (2005) and Meyerhoff et al. (forthcoming) report problems that may be encountered when including / excluding ASC in welfare estimates. It is thus safe to conclude that including / excluding ASC may be decisive for welfare estimates.

to process the information differently, and the utilisation of the name (label) altered their choices.

This paper adds to the literature by proposing a new approach to value drivers of environmental goods. In the proposed formulation estimated willingness to pay consists of two sub-components: a function of a physical attributes of the good and a value of the label, which is used for presenting the good or valuation scenario to the respondents in stated preference methods. The proposed framework provides a new insight into preferences and elicited WTP, and has an important meaning for the application of welfare estimates. Our study provides evidence that labels may be a substantial constituent of estimated value and demonstrate that this may be a reason for insufficient sensitivity to scope of welfare estimates. Therefore, accepting value of a label as one of the components of estimated value of a good provides an alternative explanation of the sources of potential problems with the scope test.

Acknowledgments

The authors wish to thank Tomasz Żylicz whose invaluable support and assistance made this study possible. The contribution by Małgorzata Buszko-Briggs is also greatly acknowledged. The research was funded by Polish Ministry of Science and Higher Education and the Foundation for Polish Science.

References

Adamowicz WL, Boxall P, Williams M, Louviere JJ (1998) Stated Preference Approaches to Measuring Passive Use Values: Choice Experiments versus Contingent Valuation. American Journal of Agricultural Economics 80: 64-75

Adamowicz WL, Boxall PC, Louviere JJ, Swait JD (1999) Stated Preference Methods for Valuing Environmental Amenities. In: Bateman IJ, Willis KG (eds) Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the USA, EC, and Developing Countries, Oxford University Press, New York

Andreoni J (1989) Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. Journal of Political Economy 97(6): 1447-1458

Andreoni J (1990) Impure Altruism and Donations to Public Goods: a Theory of Warm-glow Giving. Economic Journal 100: 464-477

Arrow K, Leamer EE (1994), Comment No. 87. submitted to NOAA in response to Advanced Notice of Proposed Rulemaking (59 FR 1062)

Bartczak A, Lindhjem H, Navrud S, Zandersen M, Żylicz T (forthcoming) Valuing Forest Recreation on the National Level in a Transition Economy: The Case of Poland. Forest Policy and Economics:

Bateman IJ, Brouwer R (2006) Consistency and construction in stated WTP for health risk reductions: A novel scope-sensitivity test. Resource and Energy Economics 28(3): 199-214

Bateman IJ, Cole M, Cooper P, Georgiou S, Hadley D, Poe GL (2004) On visible choice sets and scope sensitivity. Journal of Environmental Economics and Management 47(1): 71-93

Becker G (1974) A Theory of Social Interactions. Journal of Political Economy 82: 1095-1117

Bennett J, Rolfe J, Morrison MD (2001) Remnant Vegetation and Wetlands Protection: Non-market Valuation. In: Bennett J, Blamey RK (eds) The Choice Modelling Approach To Environmental Evaluation, Edward Elgar, Cheltenham

Bergstrom JC, Reiling SD (1998) Does Existence Value Exist? In) Multiple Objective Decision-Making for Land, Water, and Environmental Management, Boca Raton, FL, Lewis Publishers

Biénabe E, Hearne RR (2006) Public preferences for biodiversity conservation and scenic beauty within the framework of environmental services payments. Forest Policy and Economics 9: 335-348

Birol E, Karousakisb K, Koundouric P (2006) Using a Choice Experiment to Account for Preference Heterogeneity in Wetland Attributes: The Case of Cheimaditida Wetland in Greece. Ecological Economics 60: 145-156

Blamey RK, Bennett JW, Louviere JJ, Morrison MD, Rolfe J (2000) A test of policy labels in environmental choice modelling studies. Ecological Economics 32(2): 269-286

Brookshire DS, Ives BC, Schulze WD (1976) The Valuation of Aesthetic Preferences. Journal of Environmental Economics and Management 3: 325-346

Brouwer R, Langford IH, Bateman IJ, Turner RK (1999) A Meta-Analysis of Wetland Contingent Valuation Studies. Regional Environmental Change 1: 47-57

Carson RT (1997) Contingent Valuation and Tests of Insensitivity to Scope. In: Kopp R, Pommerhene W, Schwartz N (eds) Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods, Kluwer, Amsterdam

Carson RT, Mitchell RC (1993) The Issue of Scope in Contingent Valuation Studies. American Journal of Agricultural Economics 75(5): 1263-1267

Carson RT, Mitchell RC (1995) Sequencing and Nesting in Contingent Valuation Surveys. Journal of Environmental Economics and Management 28: 155-173

Carson RT, Hanemann WM (2005) Contingent Valuation. In: Mäler KG, Vincent JR (eds) Handbook of Environmental Economics, Elsevier, Amsterdam

Carson RT, Flores NE, Meade NF (2000), Contingent Valuation: Controversies and Evidence. Working paper

Chilton SM, Hutchinson WG (1999) Some Further Implications of Incorporating the Warm Glow of Giving into Welfare Measures: A Comment on the Use of Donation Mechanisms by Champ et al. Journal of Environmental Economics and Management 37(2): 202-209

Clarke KA (2003) Nonparametric Model Discrimination in International Relations. The Journal of Conflict Resolution 47(1): 72-93

Clarke KA (2007) A Simple Distribution-Free Test for Nonnested Model Selection. Political Analysis 15(3): 347-363

Cooper P, Poe GL, Bateman IJ (2004) The Structure of Motivation for Contingent Values: a Case Study of Lake Water Quality Improvement. Ecological Economics 50(1-2): 69-82

Cummings RG (1989), Comments Concerning Proposed U.S. Department of Interior Type B Natural Resource Damage Assessment Rules. submitted to U.S.D.O.I. on behalf of ASARCO, Idarado Mining Company and Newmont Mining Corporation

Diamond PA, Hausman JA (1994) Contingent Valuation: Is Some Number better than No Number? The Journal of Economic Perspectives 8(4): 45-64

Garber-Yonts BE, Kerkvliet J, Johnson R (2004) Public Values for Biodiversity Conservation Policies in the Oregon Coast Range. Forest Science 50: 589-602

Goodstein ES (1995) Economics and the Environment. Prentice Hall, Englewood Cliffs, NJ

Gunn AS (1980) Why Should We Care about Rare Species? Environmental Ethics 2: 17–37

Hanemann MW (1982), Applied Welfare Analysis with Qualitative Response Models. University of California, Berkeley

Hausman JA (1993) Contingent Valuation: A Critical Assessment. North-Holland, Amsterdam

Heberlein TA, Wilson MA, Bishop RC, Schaeffer NC (2005) Rethinking the scope test as a criterion for validity in contingent valuation. Journal of Environmental Economics and Management 50(1): 1-22

Hensher DA, Greene WH (2002) Specification and estimation of the nested logit model: alternative normalisations. Transportation Research Part B: Methodological 36(1): 1-17

Horne P (2006) Forest Owners' Acceptance of Incentive Based Policy Instruments in Forest Biodiversity Conservation – A Choice Experiment Based Approach. Silva Fennica 40(1): 169–178

Horne P, Boxall CP, Adamowicz WL (2005) Multiple-use Management of Forest Recreation Sites: A Spatially Explicit Choice Experiment. Forest Ecology and Management 207: 189-199

Jacobsen JB, Thorsen BJ (forthcoming) Where to Put a National Park and What to Put in it? An a Priori Study of the Willingness-To-Pay for Coming National Parks.
Jacobsen JB, Boiesen JH, Thorsen BJ, Strange N (2008) What's in a Name? The Use of Quantitative Measures Versus 'Iconised' Species when Valuing Biodiversity. Environmental and Resource Economics 39(3): 247-263
Johansson-Stenman O (1998) The Importance of Ethics in Environmental Economics with a Focus on Existence Values. Environmental and Resource Economics 11(3-4): 429–442
Kahneman D (1986) Comments. In: Cummings RG, Brookshire DS, Schulze WD (eds) Valuing Environmental Goods: An Assessment of the Contingent Valuation Method, Rowman and Allanheld, Totowa, NJ
Kahneman D, Knetsch JL (1992) Valuing Public Goods: The Purchase of Moral Satisfaction. Journal of Environmental Economics and Management 22: 57-70
Kahneman D, Ritov I (1993), Determinants of Stated Willingness to Pay for Public Goods: A study in the Headline Method. Department of Psychology, University of California, Berkeley
Kahneman D, Tversky A (2000) Choices, Values, and Frames. Cambridge University Press, Cambridge, UK
Knetsch JL, Sinden JA (1984) Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value. The Quarterly Journal of Economics: 507-521
Krinsky I, Robb AL (1986) On Approximating the Statistical Properties of Elasticities. The Review of Economics and Statistics 68(4): 715-719
Lancaster K (1966) A New Approach to Consumer Theory. Journal of Political Economy 84: 132-157
Lehtonen E, Kuuluvainen J, Pouta E, Rekola M, Li C-Z (2003) Non-market Benefits of Forest Conservation in Southern Finland. Environmental Science & Policy 6(3): 195-204
Louviere JJ, Hensher DA, Swait JD (2006) Stated Choice Methods: Analysis and Applications. Cambridge University Press, Cambridge
McClure SM, Li J, Tomlin D, Cypert KS, Montague LM, Montague PR (2004) Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. Neuron 44: 379-387
McConnell KE (1983) Existence and Bequest Value. In: Rowe R, Chestnut L (eds) Managing Air Quality and Scenic Resources at National Parks and Wilderness Areas, Westview Press, Boulder, CO
McConnell KE (1997) Does Altruism Undermine Existence Value. Journal of Environmental Economics and Management 32: 22-37
McFadden D (1974) Conditional Logit Analysis of Qualititative Choice Behaviour. In: Zarembka P (ed) Frontiers in Econometrics, Academic Press, New York, NY
Meyerhoff J, Dehnhardt A, Hartje V (forthcoming) Take Your Swim Suit Along… - the Value of Improving Urban Bathing Sites in the Metropolitan Area of Berlin.
Mogas J, Riera P, Bennett J (2005) Accounting for Afforestation Externalities: A Comparison of Contingent Valuation and Choice Modelling. European Environment 15: 44-58
Nielsen AB, Olsen SB, Lundhede T (2007) An Economic Valuation of the Recreational Benefits Associated with Nature-Based Forest Management Practices. Landscape and Urban Planning 80: 63-71
Olson M (1965) The Logic of Collective Action. Harvard University Press, Harvard
Park T, Loomis JB, Creel M (1991) Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies. Land Economics 67(1): 64

Poe GL, Severance-Lossin EK, Welsh MP (1994) Measuring the Difference (X - Y) of Simulated Distributions: A Convolutions Approach. American Journal of Agricultural Economics 76(4): 904-915

Poe GL, Boyle KJ, Bergrsom JC (2000) A Meta Analysis of Contingent Values for Groundwater Contamination in the United States. Paper presented at the European Association of Environmental and Resource Economists, Rythymnon, Greece

Poe GL, Giraud KL, Loomis JB (2005) Computational Methods for Measuring the Difference of Empirical Distributions. American Journal of Agricultural Economics 87(2): 353-365

Powe NA, Bateman IJ (2004) Investigating Insensitivity to Scope: A Split-Sample Test of Perceived Scheme Realism. Land Economics 80(2): 258-271

Rescher N (1980) Unpopular Essays on Technological Progress. University of Pittsburgh Press, Pittsburgh

Ritov I, Kahneman D (1997) How People Value the Environment. In: Bazerman M, Messick D, Tenbrunsel A, Wade-Benzoni K (eds) Environment, ethics, and behavior, New Lexington Press, San Franciso, CA

Roe B, Haab T (2007) Using Biomedical Technologies to Inform Economic Modeling: Challenges and Opportunities for Improving Analysis of Environmental Policies. Paper presented at the Frontiers in Environmental Economics Conference, Washington, DC, February 26-27, 2007

Rolfe J, Bennett J, Louviere J (2000) Choice Modelling and its Potential Application to Tropical Rainforest Preservation. Ecological Economics 35: 289-302

Rolston H (1988) Environmental Ethics: Duties to and Values in the Natural World. Temple University Press, Philadelphia

Rosenberger R, Loomis J, Shrestha R (1999) Meta-Analysis of Outdoor Recreational Use Value Estimates: Convergent Validity Tests. Paper presented at the Annual U.S.D.A. W-133 Meeting, Tucson, Arizona

Rossiter JR, Percy L (1997) Advertising Communications and Promotion Management. McGraw Hill, Sydney

Russow LM (1981) Why Do Species Matter? Environmental Ethics 3: 101–112

Schwartz N (1997) Cognition, Communication, and Survey Measurement: Some Implications for Contingent Valuation Surveys. In: Kopp RJ, Pommerhene WW, Schwartz N (eds) Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods, Kluwer, Boston

Schwarz G (1978) Estimating the Dimension of a Model. The Annals of Statistics 6(2): 461-464

Shiell A, Gold L (2002) Contingent Valuation in Health Care and the Persistence of Embedding Effects Without the Warm Glow. Journal of Economic Psychology 23(2): 251-262

Smith VK, Osborne LL (1996) Do Contingent Valuation Estimates Pass a “Scope” Test? A Meta-analysis. Journal of Environmental Economics and Management 31(3): 287-301

Svedsäter H (2000) Contingent valuation of global environmental resources: Test of perfect and regular embedding. Journal of Economic Psychology 21(6): 605-623

Vuong QH (1989) Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses. Econometrica 57(2): 307-333

Walsh RG, Johnson DM, McKean JR (1992) Benefit Transfer of Outdoor Recreation Demand Studies: 1968-1988. Water Resources Research 28: 707-713

Watson DO, McFarlane BL, Haener MK (2004) Human Dimensions of Biodiversity Conservation in the Interior Forests of British Columbia. BC Journal of Ecosystems and Management 4: 1-20

Xu W, Lippke BR, Perez-Garcia J (2003) Valuing Biodiversity, Aesthetics, and Job Losses Associated with Ecosystem Management Using Stated Preferences. Forest Science 49: 247-257

Figure 1. Example of a choice card

	Option A:	Option B:	Option C:
	Status Quo	Extension of the National Park	Other Form of Protection
Natural Ecological Proceses	**no change –** protection of natural ecological proceses at 16% of the forest area	**no change –** protection of natural ecological proceses at 16% of the forest area	**no change –** protection of natural ecological proceses at 16% of the forest area
Rare Species of Fauna and Flora	**no change –** decline threatening extinction	**substantial improvement –** better condition of current standings and their expansion	**partial improvement –** maintaining and better condition of current standings
Ecosystem Components	**no change –** lack of some components and decrease in quality of the existing ones	**minor improvement –** regeneration of deteriorated components on 10% of the forest area	**partial improvement –** regeneration of deteriorated components on 30% of the forest area
Cost – your tax increase (yearly)	0 zł	50 zł	10 zł
CHOICE	☐	☐	☐

Faculty of Economic Sciences
University of Warsaw
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl