University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 28/2015 (176)

Heini Ahtiainen
Janne Artell
Mikołaj Czajkowski
Jürgen Meyerhoff

Performance of Different Approaches in International Benefit Transfer: Insights from a Nine Country Experiment

Warsaw 2015

Performance of different approaches in international benefit transfer: Insights from a nine country experiment

HEINI AHTIAINEN
Natural Resources Institute Finland (Luke)
e-mail: heini.ahtiainen@luke.fi

JANNE ARTELL
Natural Resources Institute Finland (Luke)
e-mail: janne.artell@luke.fi

MIKOŁAJ CZAJKOWSKI
Faculty of Economic Sciences
University of Warsaw
e-mail: miq@wne.uw.edu.pl

JÜRGEN MEYERHOFF
Technische Universität Berlin
e-mail: juergen.meyerhoff@tu-berlin.de

Abstract

This paper investigates the performance of different approaches in international benefit transfer using data from identical and simultaneous contingent valuation surveys on marine water quality in nine European countries. The environmental good is shared by the study countries, but the countries differ substantially in their income levels and geography. We test the performance of three approaches: unit value transfer, unit value transfer with income elasticity adjustment, and function transfer. We find that at least in the case of international benefit transfers (when respondents' mean incomes vary) unit value transfer with income elasticity adjustment performs the best, both in terms of mean absolute transfer errors and minimum tolerance levels for the equivalence of welfare measures. We argue that this approach should become a standard for quick and easy benefit transfer, and can serve as a baseline for comparisons with more complicated function transfers.

Keywords:

benefit transfer, contingent valuation, international, transfer errors, water quality

JEL:

Q51, Q53, Q56

Acknowledgements

The author order is alphabetical reflecting equal contribution by the authors. The authors are grateful for the financial support provided by four Finnish Ministries (Ministry of Agriculture and Forestry, Ministry of Environment, Ministry of Transport and Communications and Ministry of Finance), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas); the Danish Strategic Research Council; the German Federal Environment Agency (UBA, Fkz 3710 25 202); and the Swedish Environmental Protection Agency. The first author gratefully acknowledges the funding from the Finnish Cultural Foundation. The third author gratefully acknowledges the support of the Polish Ministry of Science and Higher Education and the Foundation for Polish Science. We would also like to thank the BalticSTERN Secretariat and team consisting of researchers from all the nine Baltic Sea states for joint work in data collection and the willingness to share the data.

1. Introduction

Benefit transfer is an estimation method that will always be inaccurate – this statement by Kristofersson and Navrud (2005) still provides a true framework for any debate about international benefit transfer (BT). It is generally accepted that in many cases the use of BT is without alternatives, as resources for primary studies are lacking or value estimates are needed in a very short time span. Richardson et al. (2014) have additionally pointed out that demand for BT is likely to increase in the future especially in the context of ecosystem service valuation. Hence, the burning question is what kind of BT approach will result in the most accurate transfers.

Benefit transfer refers to the use of existing primary value estimates from a 'study' site to predict welfare estimates for an unstudied 'policy' site (Brouwer 2000, Wilson & Hoehn 2006). As Richardson et al. (2014) note, historically most transfers are unit value transfers, where values (e.g., mean willingness to pay) from the study site are directly transferred to the policy site. Such transfers require considerable similarity between the study and policy sites. However, even with similar site-specific characteristics, transfers generally lack full compatibility with respect to time, currency, and population characteristics, such as income.

While certain conditions under which benefit transfer performs better have been identified (Kaul et al. 2013), there are numerous additional challenges to international benefit transfer (Johnston and Rosenberger 2010). For example, some authors have raised the issue that BT across countries faces additional obstacles (e.g., Ready and Navrud 2006; Lindhjem and Navrud 2008). Among the unique challenges of international BT is the need to correct for differences in currencies (Navrud 2004) and income levels (Navrud and Ready 2007). Different currencies are typically made commensurate using either market-based (nominal) or purchasing power parity (PPP) corrected exchange rates (Ready et al. 2004). Changes in price levels over time, on the other hand, are usually taken into account using the consumer price index (Ready and Navrud 2006). Despite the additional assumptions inbuilt in such adjustments, international BT remains desirable as it increases the number of available primary studies for transfer. Further, it could be essential for less developed and transitioning countries where good quality national primary valuation studies might not be available at all (Czajkowski and Ščasný 2010).

A couple of studies have dealt with factors that are of particular interest for international BT. Czajkowski and Ščasný (2010), in their study concerning water quality improvements in Poland, the Czech Republic and Norway, found support for applying an income elasticity of willingness to pay (WTP) of 1 to adjust transferred welfare estimates in case of high income differences between countries (which is equivalent of assuming that respondents' WTP is a constant share of their incomes). Bateman et al. (2011) used a five country data to argue that transfer functions should only incorporate explanatory variables suggested by the general economic theoretic principles, rather than all variables which may increase function fit but not necessarily accuracy of transfers. Hynes et al. (2013) investigated the effect of adjusting for cultural differences on transfer errors in international BT but found that after controlling for differences in income the impact of cultural adjustments was small. Overall, as noted also by Johnston and Rosenberger (2010), results of studies investigating international BT provide a mixed message, with some assessments finding that it passes convergent validity tests (e.g., Shresta and Loomis 2003, Czajkowski and Ščasný 2010, Bateman et al. 2011), others observing significant differences (e.g., Muthke and Holm-Müller 2004, Brouwer and Bateman 2005, Lindhjem and Navrud 2008) and still other studies concluding that applying meta-analysis is promising (Johnston and Thomassin 2010, Londoño and Johnston 2012).

In this study, we provide further insights into the performance of BT approaches by transferring values between nine European countries. Building on the suggestions made by Bateman et al. (2011) and Czajkowski and Ščasný (2010), we compare the performance of three different BT approaches through examining transfer errors (TE) and the minimum tolerance levels (MTL) that make observed and transferred values statistically equivalent[1] (Kristofersson and Navrud 2005, Czajkowski and Scasny 2010). The first two BT methods in our comparisons transfer adjusted mean WTP values, first correcting for purchasing power parity only, and then, additionally, accounting for income differences using the income elasticity of WTP correction. In the third case, we employ a function transfer wherein the function comprises only of variables reflecting theoretic principles identified in earlier literature (Bateman et al. 2011).

The data come from an internationally coordinated contingent valuation (CV) study on the benefits of reaching good ecological status in the Baltic Sea – a policy target which mostly focuses on marine waters eutrophication. The survey was conducted simultaneously in all the coastal states around the sea, resulting in more than 10,500 completed responses. The dataset includes a larger number of countries and observations than previous international BT tests, and thus offers an opportunity for a more robust comparison. Moreover, the important characteristic of our study is that the valued good is identical[2] for all respondents ("water quality in the Baltic Sea"). This eliminates, at least to some extent, differences across sites that are generally a concern for BT studies. This setting, along with the

[1] i.e. the minimum difference which would result in the rejection of the null hypothesis of equivalence of two values.
[2] Identical in terms of what the survey portrayed to the respondents.

fact that the observed WTP vary considerably between the countries (Ahtiainen et al. 2014), provides excellent grounds for testing the performance of different approaches in international BT.

Overall, we find that a unit value transfer with income elasticity adjustment (using income elasticity of WTP equal to 1) performs the best of the three approaches we examined. The value function transfer, even if based only on theory-driven explanatory variables (as suggested by Bateman et al. 2011), results in higher transfer errors and equivalence levels despite using more explanatory variables. We argue that this approach, due to its simplicity and consistently very good performance, should become a standard for quick and easy international benefit transfer.

2. Previous BT studies using same survey instrument

To our knowledge, previous studies on the performance of international benefit transfers that use the same survey instrument are relatively rare.[3] Among those we identified are Barton and Mourato (2003) who assessed transferring WTP for avoided adverse health effects from water pollution between Portugal and Costa Rica, finding transfer errors in the range of 45-130%. Adjusting for income or other socio-demographic variables did not reduce TE. Ready et al. (2004) found TE between 20% and 85% in a five-country study of WTP to avoid air and water pollution health impacts, and income adjustment or value function transfer did not reduce transfer errors compared to unadjusted unit value transfer. In a BT between France and Germany, Rozan (2004) suggested using a BT approach only if errors of 30% or more are acceptable. Brouwer and Bateman (2005) examined the performance of unit and function transfer of WTP for health risk reductions transferring values between four countries. They found unadjusted unit value transfers to produce lower TE for similar countries whereas function transfers performed better for dissimilar countries. In the context of air pollution, Abou-Ali and Belhaj (2005) reported TE in the range of 60-220% between two developing countries, Egypt and Morocco. Adjusting for income and other socio-demographic factors reduced errors. Kristófersson and Navrud (2007) used identical surveys in three Nordic countries, i.e., Norway, Sweden and Iceland, to elicit use and non-use values of freshwater fish populations and subsequently transfer values between the countries. Using equivalence tests for the transferred values they observed that the WTP estimates were consistent between the three countries for all tested scenarios. At the same time they pointed

[3] As the present paper concerns an international BT based on data originating from using the same survey instrument, we confine this review to studies employing the same survey instrument in each country. Hence, studies that have used meta-analysis for international BT (e.g., Lindhjem and Navrud 2008, Hynes et al. 2013) are not considered here.

out that the accuracy relies heavily on the similarity of the populations and improvement scenarios used in each country. Finally, in the only study we are aware of that used WTP estimates from choice experiments in international benefit transfer, Kosenius and Ollikainen (2015) found that regardless of the adjustments for differences in income and standard of living, value transfer from one country to another can result in substantial errors. For transfers of WTP estimates for changes in coastal habitats at the Baltic Sea across Sweden, Finland, and Lithuania they found TE of up to 400%.

Three studies thus far have examined water quality value transfer using similar survey instruments in an international context. Muthke and Holm-Müller (2004) tested international BT between Germany and Norway for water quality improvements in lakes. Based on their results they concluded that international BT is inadvisable. Czajkowski and Ščasný (2010) investigated the validity of benefit transfers between countries that have different income levels. Based on valuation studies of lake water quality improvements they showed that income adjustment improved the accuracy of BT. Specifically, they found support for using income elasticity of WTP equal to one for transfers between countries with heterogeneous income levels. Bateman et al. (2011) examined the transfer of benefits for freshwater quality improvements between five countries. They concluded, as Brouwer and Bateman (2005) that simple unit value transfers are to be preferred when transfers involve similar sites, but function transfers lead to lower TE with dissimilar sites. Bateman et al. (2011) also argued that function transfers should be based on theory-driven models that include only variables for which there are prior expectations instead of models that are solely driven by statistically significant covariates.

This brief review of studies using the same survey instrument in different countries provides mixed evidence for BT across countries. Some studies find no effect of adjusting for socio-demographic factors while others suggest that controlling for the dissimilarities between countries, income levels in particular, either via adjustments based on assumptions with respect to constant income elasticity or function transfer can substantially improve BT accuracy. Our study provides more evidence on the issue.

3. Methods and data

3.1 Benefit transfer approaches

Generally, there are two main approaches to benefit transfer: unit value transfer and function transfer (Navrud and Ready 2007). Unit value transfer can be implemented in various ways. Its simplest form assumes equal marginal values for an average individual at the original study site and the policy site. This assumption might be too restrictive especially in an international context as the populations at the sites are likely to differ. Therefore, extensions of the simple unit transfer have been introduced. The most common adjusts for differences in mean income levels between a study and a policy site. According to Navrud and Ready (2007), such an adjustment requires two practical decisions by the researcher. The fist is to assume the income elasticity, and the other one is to choose an appropriate exchange rate to measure the values in a common currency.

There is some evidence that environmental goods might be luxury goods, implying income elasticity of demand to be higher than one (Ghalwash 2008). However, Flores and Carson (1997) have shown that the relation between income elasticity of demand and income elasticity of WTP is not straightforward, and in the case of public good valuation, knowledge of the one does not provide information on the other. Several studies provide evidence that income elasticity of WTP for environmental goods may be less than one (Kriström and Riera 1996, Ready et al. 2002, Hökby and Söderqvist 2003, Jacobsen and Hanley 2009). These studies show that income elasticity of WTP ranges usually between 0 and 1. On the other hand, Pearce (2006) and Czajkowski and Ščasný (2010) found that arbitrarily setting income elasticity of WTP to 1 performed the best when adjusting transferred welfare estimates in the presence of high income differences. They argue that assuming a constant income elasticity of 1 has a convincing interpretation – respondents declare WTP for a particular good which is a constant share of their incomes, irrespectively of what their levels are.

Another commonly applied adjustment is to account for the differences in price levels between the study and policy site using the purchasing power of currencies. Official (financial) exchange rates may reflect political and macroeconomic risk factors. The purchasing power parity (PPP) adjusted exchange rates, developed by the U.S. International Comparison Program (ICP), reflect the purchasing power of currencies in different countries or regions, and are preferred to financial exchange rates in international BT because they reflect the differences in price levels and thus lead to more meaningful international comparisons (e.g., Ready and Navrud 2006).

The second main approach for transferring benefit estimates is to use function transfer.[4] In this case the so called value function estimated using the data available at the study site is used to predict policy site value. This approach exploits more information about the differences between the study and the

[4] Another BT technique that can be used to estimate the value function, if more than one study with value estimates is available, is meta-analysis. As we do not use this approach here we guide the reader to Navrud and Ready (2007) for a comprehensive review of this technique.

policy site, as well as between the affected populations. Values at the policy site are predicted using independent variables collected from secondary data at the policy site and parameter values estimated from the study site. As Rosenberger and Loomis (2003) have stated, potential drawbacks of conducting a function transfer are that the values of the independent variables for the policy site have to be known and that this approach also implies that the statistical relationship between the dependent and the independent variables are assumed to be the same at both study and policy site. In an international context, the latter means that e.g., cultural and recreation-related factors have a similar effect on WTP across countries.

Taking these drawbacks as a point of departure, Bateman et al. (2011) suggested some guidance on the specification of appropriate value transfer functions. They noted that the functions should be developed from theoretical instead of ad-hoc statistical approaches. Accordingly, they argued that theory suggests that the benefits of an environmental change should be determined by the change in provision, site characteristics, the availability of substitutes, and characteristics of the individual common to all utility functions (e.g., the individual's income). They found that unit value transfers are likely to give defensible welfare estimates if transfers involve broadly similar provision changes of similar goods across similar contexts. If similarities fail to hold, then function transfers will result in lower transfer errors if they are constructed using general economic theoretic principles, i.e. contain only the variables for which economic theory predicts they would influence observed WTP.

3.2 Measuring the quality of benefit transfer

The observed and transferred values are likely to differ. One way to measure their relative difference is:

$$TE = \frac{WTP_{transferred} - WTP_{observed}}{WTP_{observed}}, \quad (1)$$

where *TE* stands for transfer error, which can be either positive or negative and is expressed in relation to the observed value (Kirchhoff et al. 1997). This approach, however, does not take into account the uncertainty associated with the value estimates.

The alternative was proposed by Kristofersson and Navrud (2005). They argue that instead of testing if the transferred and observed values are *equal* ($H_0 : WTP_{transferred} = WTP_{observed}$) one should test if the values are *equivalent*, given a transfer error (or a difference D) which is acceptable for a particular application $H_0 : \left| WTP_{transferred} - WTP_{observed} \right| \geq D$. Note, however, that the null hypothesis is now reversed and thus, in the case of highly uncertain estimates, statistical inference leads to different conclusions. To illustrate, imagine two equal estimates with relatively low confidence intervals (i.e. small standard errors). In such a case the non-rejection region of the equivalence hypothesis is larger than the rejection region of the equality hypothesis. On the contrary, for estimates with very large standard errors (very uncertain) it may never be possible to say that one is statistically different from the other (reject the equality hypothesis). Sufficiently high uncertainty would, on the other hand, make it impossible to conclude that the values are equivalent (reject the non-equivalence hypothesis). As a result, evaluating the validity of a benefit transfer using the Kristófersson and Navrud (2005) equivalence approach is a way to penalize uncertainty.

Czajkowski and Scasny (2010) extend this approach by proposing to estimate Minimum Tolerance Levels (MTL) for equivalence at 95% confidence level, rather than using a tolerance level specified a priori and calculating statistical certainty that the values are equicalent. Their approach searches for the minimum difference *D* (corresponding to the minimum TE) which would result in the rejection of the null hypothesis of the equivalence at the usual significance level of 5%. Formally, the MTL is a θ which solves the following equation[5]:

$$\min \theta \in [0, +\infty) \quad \text{s.t.} \quad \Pr\left(\left| WTP_{transferred} - WTP_{observed} \right| \geq \theta WTP_{observed} \right) < 0.05 \ . \qquad (2)$$

In what follows, we use both of these approaches, calculating mean absolute transfer errors (MTE) and minimum tolerance levels (MTL) to evaluate the validity of different benefit transfer techniques for our dataset.

3.3 Survey design

The benefit transfer in this article is based on an international contingent valuation (CV) study (Ahtiainen et al. 2014). The CV study elicited willingness to pay for reducing eutrophication in the Baltic Sea with national surveys conducted in all nine countries bordering the sea: Denmark, Estonia, Finland, Germany, Latvia, Lithuania, Poland, Russia and Sweden. The survey was developed in international cooperation, translated into national languages and tested with expert reviews, cognitive interviews,

[5] A custom Matlab function derived for calculating minimum tolerance levels allowing for the equivalence of 2 WTP estimates is available for download at http://czaj.org/research/estimation-packages.

focus groups and pilot surveys before the final survey. This resulted in using an identical survey instrument in each country, also in terms of the valuation scenario.

Willingness to pay was assessed for two valuation scenarios based on the improved eutrophication status of the Baltic Sea in 2050 following the delivery of (i) 100%, and (ii) 50% of the nutrient load reduction targets specified in HELCOM's Baltic Sea Action Plan (BSAP; HELCOM 2007). The improvement scenarios were contrasted with a baseline (reference) level of eutrophication in 2050 that assumed the no change in the wastewater treatment infrastructure, agricultural practices and other economic sectors. The scenarios and the baseline were developed from the predictions of state-of-the-art Baltic Sea marine models (Ahlvik et al. 2014, Kiirikki et al. 2001, 2006, Maar et al. 2011) and in cooperation with marine ecologists. The models entailed improvements in eutrophication status across the entire Baltic Sea, determining specific eutrophication levels for each of its nine sub-basins.

In the CV survey, the eutrophication levels were described using a five-step water quality ladder. The ladder included five different ecosystem effects: water clarity, blue-green algal blooms, state of underwater meadows, fish species composition and oxygen depletion at deep sea areas. After introducing the respondents to the ladder, the scenarios were shown in corresponding colour-coded maps that depicted the level of eutrophication in each sub-basin of the sea in 2050. The use of maps, instead of verbal-only description of scenarios, was considered to improve respondent comprehension (see e.g., Bateman et al. 2009 and Bateman et al. 2011). Each questionnaire included both the large (100%) and small (50%) improvement scenario in a varying order to account for possible scope and order effects. While the Baltic Sea Action Plan served as the basis for scenario development, it was not explicitly mentioned to the respondents to maintain the focus on the environmental change rather than international politics.

The payment vehicle was described as a common 'Baltic Sea tax', collected from each individual and company in all Baltic Sea countries. The survey included an in-the-market, or spike, question for being willing to pay in principle. After showing the eutrophication maps, respondents were requested to state their WTP for improving the condition of the Baltic Sea from the baseline to the policy level using country-specific payment cards. The payment cards were constructed similarly in each country based on the WTPs observed in the pilot studies, although specific bid values differed. The range was designed so that neither the lower nor the upper end of the bid distribution would be truncated as this can affect the welfare estimates (Rowe et al. 1996, Roach et al. 2002).

In our case, the environmental good in question, i.e. the Baltic Sea marine environment, is shared by all the study countries. The respondents were informed that the improvements which they were asked to value would occur in *open-sea areas* (as opposed to coastal areas) in the entire Baltic Sea, ensuring that the valued good was common to all countries.[6]

3.4 Econometric framework for calculating welfare measures

The payment card used as the elicitation format in the study can be interpreted to yield interval-type data, since respondents' true maximum WTP lies between the chosen bid and the next largest bid in the payment card (Cameron and Huppert 1989). We used this information to fit parametric distributions of respondents' WTP. We assumed that respondents' WTP followed a normal distribution. However, as WTP was expected to be non-negative and a considerable share of responses were identified as zero WTP, we used the spike model approach to account for that (Kriström 1997). In the spike model, respondents who declared they are 'not in the market', and respondents who have an implied probability of observing negative values, were cumulated in a spike-discontinuity of the WTP distribution at 0.

More formally, the probability that individual *i*'s WTP lies between a selected bid t_k and the next largest bid, t_{k+1}, conditional on being 'in the market' can be expressed as:

$$P(y_i = t_k) = P(t_k > WTP > t_{k+1}) = \Phi(t_{k+1}) - \Phi(t_k), \tag{3}$$

where the cumulative distribution function (CDF) at the unselected bid t_{k+1}, less the CDF at selected bid t_k gives the probability that respondents' WTP is in the interval described by these two bids, conditional on the parameters of the fitted parametric distribution. We estimate these parameters (μ, σ^2) by maximizing the following log-likelihood function:

$$\log L = \sum_i \log\left(\Phi(\mu, \sigma^2, t_{i,k+1}) - \Phi(\mu, \sigma^2, t_{i,k})\right). \tag{4}$$

Combining this approach with the information about market participation (*yes* denotes a binary variable for respondents who declared they are in the market), the log-likelihood function for observing the particular set of choices of the *N* individuals in the sample is given by Eq. (3).

[6] Indeed, the debriefing questions indicated that respondents focused on the entire Baltic Sea, rather than any of the specific basins (e.g., the basin closest to the respondent) and understood the results would mainly occur in the open sea areas, rather than at the coasts.

$$logL = \sum_i yes\left(\Phi(\mu, \sigma^2, t_{i,k+1}) - \Phi(\mu, \sigma^2, t_{i,k})\right) + (1 - yes)\Phi(\mu, \sigma^2, 0). \quad (5)$$

Maximizing this function produces estimates of the parameters underlying the WTP distribution, which can be used to simulate moments and quantiles of WTP distribution in each country[7] and taking the uncertainty associated with the model parameters into account allows for simulating standard errors and confidence intervals of these WTP characteristics.

Finally, we note that the underlying parameters (μ, σ^2) of each distribution can incorporate observed heterogeneity of respondents by accepting respondent-specific covariates. This approach allows us to estimate function transfers which utilize information on respondents' income or distance to the site, as suggested by Bateman et al. (2011).

In summary, our approach for modelling WTP allows for (1) taking the interval nature of the payment card responses into account, (2) accounting for respondents 'not in the market', who do not fit in the continuous distribution of WTP in the remainder of the population, (3) inferences about mean WTP in the population as well as its standard deviation, other moments and quantiles, and (4) simulating standard errors associated with each of these measures, which allows for evaluating the uncertainty associated with the estimates. We found that in the case of our dataset this approach outperformed other methods of modelling respondents' WTP.

4. Results

4.1 Descriptive statistics and country-specific models

The CV survey was implemented in the fall 2011 with internet panels, except in Latvia, Lithuania and Russia where face-to-face interviews were used to ensure a representative national sample.[8] Table 1

---

[7] Note that even though we use normal as an underlying parametric distribution, the predicted WTPs which are less than 0 are censored to 0. Because of this, and because of a contribution of respondents' who are not in the market, the WTP density function has a jump discontinuity at 0. As a result, the characteristics of the predicted WTP distribution in the sample are not the same as those of the underlying normal distribution, used for modeling purposes only.

[8] The surveys in Poland used both web-based and personal interviews, allowing for a possibility to control for the potential differences driven by the survey mode.

summarizes the descriptive statistics for each country. The survey setup is described in detail in Ahtiainen et al. (2012, 2014).

As Table 1 shows, the coastal countries differ substantially with regard to mean income levels and distance to the Baltic Sea coast. The countries can be classified into two groups according to average income levels: Estonia, Latvia, Lithuania, Poland and Russia comprise the lower income group, while Denmark, Finland, Germany and Sweden are in the higher income group. The variation in the mean distance reflects differences in the land area and in the length of coastline that both vary substantially between the countries. For example, Denmark is a small country with a long coastline, and therefore has the shortest mean distance to the coast. Germany and Poland are rather similar with large land areas and short coastline, while Russia is the largest country in the world with only a short stretch of coast in the Baltic Sea. Both Finland and Sweden have long coastlines, explaining the short mean distance to the coast in these relatively large countries.

Table 1. Descriptive statistics of the sample by country

Country	Survey mode	Number of observations	Mean monthly net income in 1000 € PPP	Mean distance to the Baltic Sea coast in 100 km
Denmark	Internet panel	1,061	1.66	0.13
Estonia	Internet panel	505	0.86	0.29
Finland	Internet panel	1,645	1.59	0.60
Germany	Internet panel	1,495	1.56	3.58
Latvia	Face-to-face interviews	701	0.48	0.54
Lithuania	Face-to-face interviews	617	0.34	1.86
Poland	Face-to-face interviews, internet panel	1,023 1,006	0.84	2.96
Russia	Face-to-face interviews	1,508	0.67	8.81
Sweden	Internet panel	1,003	1.43	0.40

Country-specific WTP models (Table 2) were estimated separately for the two scopes of eutrophication reduction (half and full) using the spike modelling approach detailed in section 3.3. The dependent variable was the (censored) WTP for the program in question. All models included two explanatory variables: income and distance to the Baltic Sea coast. The results indicate that distance is a significant determinant of WTP in most countries, i.e. Denmark, Finland, Germany, Latvia, Russia and Sweden, while higher income increases WTP in all countries.

Table 2. Results from the country-specific spike models – the parameters of the underlying normal distributions

	Scope: half (50%)				**Scope: full (100%)**			
Country	Mean			Standard deviation (standard error)	Mean			Standard deviation (standard error)
	Constant (standard error)	Distance (standard error)	Income (standard error)		Constant (standard error)	Distance (standard error)	Income (standard error)	
Denmark	-15.46**	-43.67**	12.93***	84.75***	-16.93**	-47.09**	13.69***	90.45***
	(7.79)	(22.12)	(3.94)	(1.60)	(8.80)	(23.31)	(4.45)	(1.61)
Estonia	-7.78	1.71	7.79	57.46***	-15.47**	2.65	17.55***	74.58***
	(6.07)	(6.25)	(4.92)	(1.26)	(6.89)	(8.08)	(5.04)	(1.91)
Finland	4.27	-12.39***	9.87***	74.53***	-10.35*	-9.67***	20.05***	88.92***
	(5.58)	(3.05)	(2.70)	(1.34)	(6.04)	(3.52)	(2.82)	(1.31)
Germany	0.71	-1.63**	4.39***	47.9***	-1.77	-2.18*	6.62***	72.71***
	(3.93)	(0.79)	(1.53)	(0.77)	(6.74)	(1.34)	(2.66)	(1.50)
Latvia	-4.09***	-3.63***	9.22***	13.82***	-5.90***	-4.04***	12.29***	17.17***
	(1.36)	(1.18)	(1.77)	(0.35)	(1.65)	(1.44)	(2.11)	(0.38)
Lithuania	-6.14**	-0.20	16.62***	20.31***	-8.39**	-1.11	27.22***	27.89***
	(2.72)	(0.94)	(5.05)	(0.55)	(3.95)	(1.37)	(7.49)	(0.90)
Poland	-7.21***	-0.59	10.32***	28.29***	-10.36***	-0.57	14.53***	33.93***
	(1.87)	(0.42)	(1.27)	(0.28)	(2.19)	(0.50)	(1.37)	(0.41)
Russia	-36.55***	-0.54***	19.63***	48.58***	-39.74***	-0.66***	23.68***	53.66***
	(4.34)	(0.12)	(5.21)	(1.10)	(4.26)	(0.12)	(4.91)	(0.88)
Sweden	-8.61	-24.25***	43.74***	126.34***	2.27	-39.31***	47.07***	166.97***
	(18.40)	(9.84)	(11.06)	(2.62)	(17.31)	(13.98)	(10.8)	(3.47)
Variables significant at the *** 1%, ** 5% and *10% level.								

Table 3 presents country-specific mean and median WTP, together with the probability of WTP = 0 (spike prob.). There is substantial variation in the mean WTP and the probability of being willing to pay between countries. Mean WTP is highest in high income countries, i.e. Sweden, Finland and Denmark, and lowest in the relatively lower income countries; Russia, Latvia and Lithuania. Even though Estonia is in the lower income group, its mean WTP is higher than the German counterpart. A possible explanation may be that Germany has more substitutes to the Baltic Sea. The substantially higher WTP in Sweden compared to other countries has also been observed in other studies of marine water quality improvements (e.g., Kosenius & Ollikainen 2015), and Swedes are the most enthusiastic users of the Baltic Sea (Ahtiainen et al. 2013).

Table 3. Willingness to pay results, in PPP corrected EUR (2011)

	Scope: half (50%)				Scope: full (100%)			
Country	Mean (standard error)	Median (standard error)	Standard deviation (standard error)	Spike prob. 0-WTP (standard error)	Mean (standard error)	Median (standard error)	Standard deviation (standard error)	Spike prob. 0-WTP (standard error)
Denmark	34.01	1.55	49.6	0.50	35.96	1.38	52.69	0.50
	(1.61)	(2.11)	(1.27)	(0.02)	(1.69)	(2.11)	(1.28)	(0.02)
Estonia	22.66	1.18	33.32	0.50	29.98	2.03	43.65	0.50
	(1.56)	(1.91)	(1.01)	(0.03)	(1.99)	(2.82)	(1.32)	(0.03)
Finland	36.38	12.44	47.68	0.43	43.85	15.61	57.16	0.43
	(1.35)	(2.58)	(0.99)	(0.01)	(1.56)	(3.04)	(1.05)	(0.01)
Germany	20.00	1.88	28.55	0.49	29.41	1.49	42.71	0.50
	(0.77)	(1.50)	(0.57)	(0.01)	(1.24)	(1.80)	(1.06)	(0.01)
Latvia	4.77	0.00	7.52	0.55	5.84	0.00	9.29	0.55
	(0.29)	(0.05)	(0.24)	(0.02)	(0.36)	(0.04)	(0.29)	(0.02)
Lithuania	7.71	0.17	11.57	0.52	10.58	0.22	15.89	0.52
	(0.47)	(0.41)	(0.34)	(0.02)	(0.66)	(0.55)	(0.53)	(0.02)
Poland	11.16	0.23	16.42	0.50	13.64	0.54	19.87	0.50
	(0.41)	(0.41)	(0.32)	(0.01)	(0.47)	(0.70)	(0.35)	(0.01)
Russia	8.47	0.00	18.78	0.72	9.75	0.00	21.20	0.71
	(0.54)	(0.00)	(0.73)	(0.01)	(0.56)	(0.00)	(0.69)	(0.01)
Sweden	75.61	44.20	88.26	0.36	97.02	53.85	115.21	0.37
	(3.36)	(5.72)	(1.93)	(0.02)	(4.33)	(7.38)	(2.62)	(0.02)

4.2 Benefit transfer results

Mean benefits reported in table 3 were transferred using three techniques: 1) unit value transfer with purchasing power parity (PPP) adjustment, 2) unit value transfer with income adjustment, using income elasticity of 1, and 3) benefit function transfer. We conducted benefit transfers from country *i* country *j* (i.e. one country acted as the study site and the other as the policy sites) and from *n-i* countries to country *i* (i.e. eight countries acted as the study site and one as the policy site). Both scopes of the environmental change were included.[9]

Mean absolute transfer errors and minimum tolerance levels for transfers from each country to the other eight countries are presented in Table 4. The mean absolute TE range from 71% to 164%. However, the variation in transfer errors is considerable depending on the source and target country, as can be seen from the minimum and maximum errors. The income elasticity adjustment produces the lowest mean TE, outperforming the unit transfer with the PPP adjustment and function transfer.

[9] Transfers were performed only between similar scopes, i.e. small-to-small and large-to-large.

Investigation of the minimum tolerance levels for equivalence reveals an even larger difference between the performances of the income elasticity adjusted and function transfers. The equivalence test shows that in function transfer, the acceptable transfer error that would result in the rejection of the equivalence hypothesis is 289%, compared to 86% in the unit value transfer with income elasticity adjustment.

Turning to consider the results for transfers from eight countries to one country (Table 5), the general results are similar. Again, average TE are the smallest for the income elasticity adjusted transfer, although the difference between income adjusted and function transfer is smaller than for the from-one-country-to-another transfers. Again, the minimum tolerance levels show more clearly the difference between the income elasticity and function transfer approaches.

The mean TE observed in this study (44-164%) are of similar magnitude in comparison with previous international studies of water quality values that used similar surveys. Depending on the BT approach, Bateman et al. (2011) found mean TE in the range of 40-121% and Czajkowski and Ščasný (2010) in the range of 26-366% for the full set of countries.

Table 4. Mean absolute transfer errors and minimum tolerance levels (minimum and maximum in parenthesis), from country *i* to country *j* (144 transfers)

BT method	**Mean transfer errors (%)**	**Minimum tolerance levels (%)***
PPP adjusted	164	183
	(2-1561)	(15-1685)
Income elasticity = 1	71	86
	(1-460)	(12-503)
Function transfer	102	289
	(0.2-1077)	(169-1348)

***** Minimum tolerance level is the minimum difference between the transferred and observed estimate which would result in the rejection of the null hypothesis of equivalence at the 5% level.

Table 5. Mean absolute transfer errors (%) and minimum tolerance levels (minimum and maximum in parenthesis), from countries *n-i* to country *i* (18 transfers)

	Mean transfer errors (%)	**Minimum tolerance levels (%)***
PPP adjusted	120	142
	(9-444)	(24-490)
Income elasticity = 1	44	59
	(1-121)	(11-141)
Function transfer	49	80
	(0-132)	(11-205)

* Minimum tolerance level is the minimum difference between the transferred and observed estimate which would result in the rejection of the null hypothesis of equivalence at the 5% level.

5. Discussion and conclusions

Our study examines the performance of international benefit transfer by comparing transfer errors and equivalence levels for different BT approaches. The data originated from identical and simultaneous contingent valuation surveys conducted in nine European countries around the Baltic Sea in 2011. As the environmental good in question, the Baltic Sea water quality, is shared by the study countries, we are able to control for the environmental good and the change in its provision so that they are as similar as possible across countries. As in all countries a majority of the respondents cared about the water quality improvement in the entire Baltic Sea, despite differences in demographics, income and geographical factors (e.g., land area and length of coastline) which affect the mean distance to the environmental good.

A notable difference between our case and earlier benefit transfers is the scale of the environmental good and the relevant population. Water quality in the Baltic Sea has national importance to the coastal countries and non-use values form a significant portion of the total value, making a national level assessment relevant. Heterogeneity in the average distance to the Baltic Sea between countries, stemming from differences in land areas and the length of coastline, may result in larger transfer errors. The earlier international BT applications have conducted transfers on a local or regional scale, where the average distances from the study and policy site are less likely affected by geography.

Another issue that is influenced by the national importance of the environmental good is the treatment of substitutes. Our value functions differ from the theory-driven model by Bateman et al. (2011) in so far as we have not included variables describing the availability of substitutes. Substitutes for unique and iconic environmental goods, such as the Baltic Sea, are difficult to identify objectively from secondary data, rendering them less useful in practical benefit transfers and policy analysis. It may also be the case that substitutes for iconic resources do not exist, at least for non-users (Rolfe and Windle 2013), which suggests that they may be less important in BT models of nationally important environmental goods.

Overall, we find that a unit value transfer with income adjustment (using income elasticity of WTP equal to 1) performs the best of the three approaches we examined. This is confirmed by the lowest mean transfer errors and minimum tolerance levels for the equivalence of welfare measures both in one-to-one and many-to-one transfers. The value function transfer, even if based only on theory-driven explanatory variables (income and distance; Bateman et al. 2011), results in higher transfer errors and equivalence levels.

A possible explanation why the unit value transfer outperforms the value function transfer despite distance is taken into account can be related to the functional form of income. Majority of function transfers assume a particular (linear and additive) relationship between WTP and income. Although our income elasticity approach assumes that the proportion of WTP to income is constant in each country, it is essentially a non-linear (log-log) functional form, which appears to perform better than income entering the WTP function additively and linearly. This offers a conclusion for future function transfer approaches, to investigate various functional forms for the value function, rather than use an ad hoc assumption which can be outperformed by even the simple unit value transfers in some cases.

Our results support the conclusions of Czajkowski and Ščasný (2010) who proposed using the income elasticity approach to control for income differences in transfers between countries. Despite the fact that the countries in our study are very heterogeneous with respect to average income, our findings indicate that the income elasticity adjusted unit value transfer outperforms PPP adjusted and function transfer.

As this point, two cautionary notes are in order. First, there are other specifications one could use for the function transfer, and in particular other transformations of income than having it enter WTP function linearly. We note that this issue has not received a lot of attention in the literature so far and our results indicate that the functional form plays a significant role for TE and MTE. Second empirical results show that the income elasticity of WTP is likely non-constant within each country (Czajkowski and Ščasný 2010, Barbier, Czajkowski, and Hanley forthcoming). It is not clear if one can expect income elasticity of WTP to change as everyone's income levels change or do income elasticities of WTP of consumers in different income groups remain relatively constant. In addition, we are not aware of any theoretical and methodological framework which would allow to incorporate non-linearities in the income elasticity of WTP and use them for predicting WTP changes as respondents' income levels change. It remains an open question and an interesting topic for future studies to establish if these other approaches (1) provide a significantly better fit to existing data and (2) allow for lower transfer errors or minimum equivalence levels.

What our study does show, however, is that using a constant income elasticity adjustment, while very simple to implement (an advantage which would be particularly valuable for a 'quick and easy' valuation method such as a benefit transfer) consistently leads to very good performance in terms of low TE and MTL. We believe this approach can become a new standard for quick and easy international benefit transfer and should be used as a baseline for comparisons with more complicated function transfers. Further studies with a similar design, i.e., same questionnaires and similar environmental changes, such as the one presented by Bateman et al. (2011) or the one presented here, are necessary to shed more light on the performance of international benefit transfer.

References

Abou-Ali H, Belhaj M (2005). Does benefit transfer always work: A multi-country comparison. Working paper in Economics No.158, Department of Economics, Gothenburg University

Ahlvik L., Ekholm P, Hyytiäinen K, Pitkänen H (2014) An economic-ecological model to evaluate impacts of nutrient abatement in the Baltic Sea. Environmental Modelling & Software 55: 164-175

Ahtiainen H, Artell J, Czajkowski M et al (2014) Benefits of meeting nutrient reduction targets for the Baltic Sea – results from a contingent valuation study in the nine coastal states. Journal of Environmental Economics and Policy 3, 278-305

Ahtiainen H, Hasselström L., Artell J et al. (2012) Benefits of meeting the Baltic Sea nutrient reduction targets - Combining ecological modelling and contingent valuation in the nine littoral states. MTT Discussion Papers 1/2012. Available at: http://www.mtt.fi/dp/DP2012_1.pdf

Ahtiainen H, Artell J, Czajkowski M et al (2013) Public preferences regarding use and condition of the Baltic Sea—An international comparison informing marine policy. Marine Policy 42: 20-30

Barbier E B, Czajkowski M, Hanley N (forthcoming) Is the income elasticity of the willingness to pay for pollution control constant?

Barton D N, Mourato S (2003) Transferring the benefits of avoided health effects from water pollution between Portugal and Costa Rica. Environment and Development Economics 8(2): 351-371

Bateman I J, Brouwer R, Ferrini S et al (2011) Making Benefit Transfers Work: Deriving and Testing Principles for Value Transfers for Similar and Dissimilar Sites Using a Case Study of the Non-Market Benefits of Water Quality Improvements Across Europe. Environmental and Resource Economics 50: 365-387

Bateman I J, Day BH, Jones AP, Jude S (2009) Reducing gains/loss asymmetry: a virtual reality choice experiment (VRCE) valuing land use change. Journal of Environmental Economics and Management 58: 106-118

Brouwer, R. 2000. Environmental value transfer: state of the art and future prospects. Ecological Economics 32:137-152

Brouwer R, Bateman I J (2005) Benefits transfer of willingness to pay estimates and functions for health-risk reductions: a cross-country study. Journal of Health Economics 24: 591-611

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: 230-246

Czajkowski M, Ščasný M (2010) Study on benefit transfer in an international setting. How to improve welfare estimates in the case of the countries' income heterogeneity? Ecological Economics 69: 2409-2416

Flores, N. E. and Carson, R. T. 1997. The relationship between the income elasticities of demand and willingness to pay. Journal of Environmental Economics and Management 33:287-295

Ghalwash, T. 2008. Demand for environmental quality: An empirical analysis of consumer behaviour in Sweden. Environmental and Resource Economics 41(1): 71-87

Hynes S, Norton D, Hanley N (2013) Adjusting for cultural differences in international benefit transfer. Environmental and Resource Economics 56:499-519

Hökby S, Söderqvist T (2003) Elasticities of Demand and Willingness to Pay for Environmental Services in Sweden. Environmental and Resource Economics 26: 361–383

Jacobsen J B, Hanley N (2009) Are there income effects on global willingness to pay for biodiversity conservation? Environmental and Resource Economics 43:137-160

Johnston, R J, Rosenberger R S (2010) Methods, trends and controversies in contemporary benefit transfer. Journal of Economic Surveys 24: 479-510

Kaul S, Boyle K J, Kuminoff N V et al (2013) What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. Journal of Environmental Economics and Management 66, 90-104

Kiirikki M, Inkala A, Kuosa H Et al (2001) Evaluating the effects of nutrient load reductions on the biomass of toxic nitrogen-fixing cyanobacteria in the Gulf of Finland, Baltic Sea. Boreal Environment Research 6: 1–16

Kiirikki M, Lehtoranta J, Inkala A et al (2006) A simple sediment process description suitable for 3D-ecosystem modeling – Development and testing in the Gulf of Finland. Journal of Marine Systems 61: 55–66

Kirchhoff S, Colby B G, LaFrance J T (1997) Evaluating the performance of benefit transfer: an empirical inquiry. Journal of Environmental Economics and Management 33:75-93

Kosenius A-K, Ollikainen M (2015) Ecosystem Benefits from Coastal Habitats – a Three-Country Choice Experiment in WTP Space. Marine Policy 58:15-27

Kristófersson D, Navrud S (2005) Validity Tests of Benefit Transfer – Are We Performing the Wrong Tests? Environmental and Resource Economics 30: 279-286

Kristófersson, D., & Navrud, S. (2007). Can use and non-use values be transferred across countries? In navrud S, Ready R (eds), Environmental value transfer: issues and methods, Dordrecht, pp. 207-225

Kriström B (1997) Spike Models in Contingent Valuation. American Journal of Agricultural Economics 79: 1013–1023

Kriström B, Riera P (1996) Is the Income Elasticity of Environmental Improvements Less Than One? Environmental and Resource Economics 7: 45–55

Lindhjem H, Navrud S (2008) How Reliable are Meta-Analyses for International Benefit Transfers? Ecological Economics 66: 425–435

Maar M., Møller EF, Larsen J et al (2011) Ecosystem modeling across a salinity gradient from the North Sea to the Baltic Sea. Ecological Modeling 222: 1696–1711

Muthke T, Holm-Mueller K (2004) National and international benefit transfer testing with a rigorous test procedure. Environmental and Resource Economics 29: 323-336

Navrud S (2004) Value transfer and environmental policy. In Tietenberg, T. and Folmer, H. (eds.) 2004. The International Yearbook of Environmental and Resource Economics 2004/2005. A survey of Current Issues. p. 189-217. Cheltenham

Pearce D W (2006) Framework for assessing the distribution of environmental quality, in Serret Y, Johnstone N (eds), The distributional effects of environmental policy, Cheltenham, pp. 23-78

Ready R, Navrud S, Day Bet al (2004) Benefit Transfer in Europe: How Reliable Are Transfers between Countries? Environmental and Resource Economics 29:67-82

Ready R, Navrud S (2006) International benefit transfer: Methods and validity tests. Ecological Economics 60: 429-434

Ready, R. C., Malzubris, J. and Senkane, S. 2002. The relationship between environmental values and income in a transition economy: surface water quality in Latvia. Environment and Development Economics 1: 147–156

Richardson L., Loomis J, Kroeger T, Casey F (2014) The role of benefit transfer in ecosystem service valuation. Ecological Economics 115:51-58

Roach B, Boyle K J, Welsh M P (2002) Testing Bid Design Effects in Multiple Bounded Contingent Valuation. Land Economics 78: 121-131

Rosenberger R S, Loomis J B (2003) Benefit Transfer, in Champ P A, Boyle KJ, Brown TC (eds.)A Primer on Non-Market Valuation. Dordrecht, pp. 445-482

Rowe RD, Schulze WD, Breffle WS (1996) A Test for Payment Card Biases. Journal of Environmental Economics and Management 31: 178–185

Rozan A (2004) Benefit transfer: a comparison of WTP for air quality between France and Germany. Environmental and Resource Economics 29: 295-306.

Wilson M A, Hoehn J P (2006) Valuing environmental goods and services using benefit transfer: The state-of-the art and science. Ecological Economics 60: 335-342

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