Socio-Economic Planning Sciences 74 (2021) 100935

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: http://www.elsevier.com/locate/seps

A Statistical Matching approach to reproduce the heterogeneity of willingness to pay in benefit transfer

Riccardo D'Alberto [a,*], Matteo Zavalloni [b], Meri Raggi [a], Davide Viaggi [b]

[a] *Department of Statistical Sciences "P. Fortunati", University of Bologna, Via Delle Belle Arti 41, 40126, Bologna, BO, Italy*
[b] *Department of Agricultural and Food Sciences, University of Bologna, Viale Fanin 50, 40127, Bologna, BO, Italy*

ARTICLE INFO

Keywords:
Benefit transfer
Willingness to pay
Statistical matching
Public goods
Agri-environmental policies

ABSTRACT

Researchers and policymakers seek a better understanding of the social demand for agri-environmental public goods (PGs), that, being nonmarket goods, are usually valuated by means of stated preference methods by eliciting people's willingness to pay (WTP). In actual policy design, benefit transfer (BT) is often preferred to novel surveys which are expensive and time demanding. Common BT approaches are value and function transfer that can provide good estimates of the mean WTP but disregard the heterogeneity of the individuals' preferences. The WTP distribution is thus flattened, leading to a misrepresentation of the PG demand. The objective of this paper is to improve BT in its ability to reproduce the actual WTP distribution at the policy site by means of the non-parametric micro Statistical Matching. We use this novel approach to transfer individual WTP values for soil erosion and carbon sequestration elicited by contingent valuation on people living in Emilia-Romagna, Italy. Comparing the results with the ones of value and function transfers, our approach outperforms the others, reflecting the actual WTP distribution and lowering the benefit transfer errors. In this way, BT can better support policymakers in designing new agri-environmental policy instruments, more targeted towards specific demand segments and hence with higher cost-effectiveness.

1. Introduction

European Union (EU) policymakers envision a re-thinking of the formulation of the agri-environmental measures of the Common Agricultural Policy (CAP). This process goes, above all, in the direction of a stronger focus on specific instruments, such as result-based, collective and value chain incentives, requiring a better understanding of the social demand for public goods (PGs).

Since the 90s, the CAP has allocated large parts of its funds to the provision of agri-environmental-climate PGs, with an estimated 30% of the CAP Rural Development Programs budget allocated to agri-environmental-climate payments in the programming period 2014–2020. However, despite this financial effort there is still a strong debate about the actual effectiveness and efficiency of these measures [1–3]. Their traditional design is a flat payment to farmers committing to given environmental-friendly practices, where the amount of the payment is supposed to cover the (average) sum of additional cost, forgone income and transaction costs resulting from the implementation of these practices. The proposal for the 2021–2027 CAP envisions a greater role for the demand side of the PGs (i.e. the societal value) as well as a more tailored design, adapting the common framework to locally heterogeneous needs and considering several policy instruments, among which result-based and/or collective payments, new market-related instruments, etc. [4]. This will finally result in a shift from compliance to performance in the CAP design, with an expected remarkable improvement in the effectiveness and efficiency of public expenditure.

In economic terms, a more accurate design of all these instruments requires a more analytical understanding of costs and benefits, by explicitly addressing the differentiation of benefits across the beneficiary population and the differentiation of costs across farms. On the one hand, this would allow a clearer identification of the cases with higher net benefit from public expenditure. On the other hand, this would help to better determine the incentives to active participation of different (segments of) actors, also considering that these measures are mostly voluntary. Focusing on the demand side, this implies a better assessment of the society's willingness to pay (WTP) for PGs and its differentiation across different individuals or groups as well as different areas.

* Corresponding author.
E-mail addresses: riccardo.dalberto@unibo.it (R. D'Alberto), matteo.zavalloni@unibo.it (M. Zavalloni), meri.raggi@unibo.it (M. Raggi), davide.viaggi@unibo.it (D. Viaggi).

https://doi.org/10.1016/j.seps.2020.100935
Received 16 December 2019; Received in revised form 10 August 2020; Accepted 10 August 2020
Available online 9 September 2020

The main challenge in assessing the demand for PGs is the lack of market prices for most of them due to their peculiar nature of nonmarket goods [5]. This means that PGs are almost never part of formal markets and their provision occurs outside these latter [6] fostering the development of the stated preference methods to quantify (i.e. estimate) a price for the PGs provision. Stated preference methods elicit the WTP through either contingent valuation [7] or choice experiments [8]. In the context of agri-environmental studies, these methods have been extensively used [9] to evaluate e.g. water quality [10], biodiversity [11], landscape features [12], and other PGs.

Considering that primary evaluations are expensive and time-consuming, in the last two decades there has been an increasing tendency to resort to use benefit transfer (BT) [13] rather than designing and fulfilling ex-novo experiments to elicit the WTP. Indeed, BT allows to transfer the WTP from existing studies (study sites) to the policy site of interest. According to Ref. [14], the two most applied BT approaches are value and function transfer. The former uses a measure of central tendency of the WTP distribution such as the mean or the median WTP, transferring it by eventually adjusting for the policy site characteristics. Function transfer is based on the estimation of the WTP function at the study site and on the subsequent transfer of the estimated coefficients at the policy site where the WTP is predicted using independent variables from secondary data. There is a third approach based on meta-analysis that lacks of a well-recognized validity [15,16] being relatively unexplored [6]. Moreover, Ref. [6] stress that the BT literature suggests to apply meta-analysis only for sub-national transfers due to the fact that transfers at higher levels should be considered unreliable. The focus of this paper is on the first two approaches.

The relative validity of value or function transfer depends on several aspects with respect to which the literature on the topic is not conclusive [17,18]. First, the transfer validity depends on the consistency and accuracy of the original WTP estimates. In this prospect the use of choice experiments or contingent valuation matters, with differences in the elicitation methods resulting in differences in the estimates validity [19]. Choice experiments are commonly believed to better include in the PGs valuation both the marginal changes in the good and the site characteristics and thus should be preferred for BT [20]. Nevertheless, practitioners use contingent valuation more than choice experiments due to its simpler applicability and broader scope [21]. In addition, the most recent literature in stated preference methods promotes the use of contingent valuation [22]. This is due to the fact that: 1) the structure of the choice options is often complex and very study site-specific, thus hardly or not easily qualified for BT and, 2) contingent valuation is often more robust and accurate in terms of the elicited WTP. Transferring WTP values elicited by contingent valuation also tends to result in lower transfer errors compared to the values transferred from choice experiments [23,24]. Second, the relative validity of the transfer depends on how much the utility function is adapted to the policy site characteristics. Then, function transfer should be better equipped to represent the heterogeneity of the individuals' preferences than value transfer. However, validity and accuracy of function transfer are often unsatisfactory and this approach can produce extremely large transfer errors (up to 7500%) [25]. Finally, how to properly reproduce the preferences heterogeneity at the policy site is still an unsolved BT challenge: choosing a functional form often leads to misrepresent the WTP distribution [13].

Considering that BT addresses a specific problem of missing information [25], we propose to use data integration methods to tackle this problem. To the best of our knowledge, no data integration method has never been applied to BT and we propose to use Statistical Matching (SM) that is one of the most recently developed methods for data integration [26]. SM is characterized by either a parametric (model-based) or a non-parametric approach. This latter outperforms parametric SM (e.g. regression and log-linear methods) in predicting the missing information [27]. Its main advantage is that it uses only "live" (or "real") observed values allowing to avoid the issues related to the *a priori* definition of any functional forms. In other words, by applying the non-parametric SM we do not have to specify any family distribution and/or estimate any model parameter. The missing values are substituted only by resorting to the available observed information such that any model mis-specification deriving from e.g. the assumption of linearity can be avoided.

We propose the use of non-parametric micro SM for benefit transfer purposes: the Statistical Matching benefit transfer approach (SMBT). SMBT serves the main objective of this paper: to improve BT in its ability to reproduce the heterogeneity of the individuals' preferences at the policy site and thus to properly reproduce the true WTP distribution. We validate SMBT and compare it with both value and function transfer in accordance with BT literature [25]. We transfer both the mean and the median WTP with value transfer while we apply the function transfer with both linear and Tobit model specifications.

We apply all these BT approaches to the WTPs elicited for a local (soil erosion) and a global (carbon sequestration) agri-environmental PGs. The elicitation method used to estimate the WTP is contingent valuation, applied by means of an on-line survey to the people living in Emilia-Romagna region (Italy). We consider two different applications: in Application 1, the province of Bologna is the policy site while the rest of the Emilia-Romagna territory serves as study site. In Application 2, the province of Ferrara is the policy site while the rest of the provinces of Emilia-Romagna are the study site. These two applications are significant examples of the transfer results sensitivity to the degree of similarity between the study site and the policy site when traditional BT approaches are used. This common topic is debated by the BT literature that at various level agrees on the fact that the greater is the correspondence between the study and the policy sites, the smaller is the expected benefit transfer error [14]. We offer further insights into the fact that a close similarity between the sites is a necessary (but not sufficient) condition to obtain WTP estimates that are robust and consistent as well as accurate and valid. We stress that there are relevant differences in the way that the classic BT approaches and the proposed SMBT perform with respect to this degree of similarity. Indeed, SMBT performs well also when the differences between the sites increase while value and function transfer do not. Moreover, only SMBT allows to reproduce the whole WTP distribution at the policy site while value and function transfers can provide, at most, good estimates of the mean/median WTP (or individual values clustered around the distribution mean).

The proposal of such a novel approach to BT allows to: 1) preserve the heterogeneity of the actual individuals' preferences at the policy site, 2) reproduce the actual WTP distribution as a whole and, 3) lower the benefit transfer error. With respect to the SM literature, this paper also enlarges the scope of the applications for such a method. Indeed, the non-parametric micro SM counts so far applications to household surveys and official administrative registers (see e.g. Refs. [28,29] and the references therein), Farm Accountancy Data Network (FADN) and Farm Structure Survey data referred to Italian farms [30] and Swiss farms [31] and FADN data and ad hoc surveys [32].

BT is usually carried out at an aggregate level with the mean/median WTP considered as a natural choice for the transfer. However, new avenues for BT applications do exist, motivated by several considerations. First, misrepresenting the WTP distribution or disregarding its relevance by transferring a (yet meaningful) WTP value such as the mean or the median study site WTP is not simply a cosmetic problem. It has instead really important policy implications. For example, the transfer of an average WTP value can distort the reconstructed demand for PGs at the policy site. This is problematic especially if we consider the relevance of the WTP distribution tails, i.e. the sub-groups of individuals that have low or high WTPs. Not considering the whole WTP distribution and thus distorting the PGs demand at the policy site can lead policymakers to substantial errors in design and ranking priorities. This, in turn, leads to budget misallocation or mismatches in spatial targeting and it is even more troublesome whenever the policy measures, rather than targeting the whole population are explicitly focused on small numbers of target beneficiaries that e.g. correspond to the higher WTP tail of the demand function. Second, from a practical point of view, data shortage is not necessarily an issue anymore. Indeed, both the increasing big data flood and the measures adopted by public institutions are progressively transforming such an issue in an addressable challenge. On the one hand, in the agri-environmental research field there has been a shift towards a big data application system that is fostering a massive production of several kinds of (raw) data that can be used for decision-making and many other purposes [33]. On the other hand, official statistics is envisaging a system of recursive data integration that puts together the information from official and ad hoc surveys, research data, administrative registers, big data, etc. [34]. In addition, the EU with its research and innovation programme Horizon 2020 is promoting the production and use of open-access, re-useable, primary data sources. These criteria are acquiring importance such that the research funding is going to be bonded to them in the upcoming future [35]. These information opportunities do open to unexplored paths and, in such a context, SMBT can be profitably used in order to address also disaggregated exercises of BT. Clearly, these considerations are not limited to the agri-environmental research field but they involve all the cases where the elicitation of nonmarket goods is carried out by means of BT.

The paper is structured as follows: in section 2 we review the state of the art in BT, discussing the actual approaches of value and function transfer, their pending challenges and the proposed SMBT. In section 3 we present the results of Application 1 and Application 2, evaluating the benefit transfer errors produced and comparing the performances of the different methods. In section 4 we discuss the results while in section 5 we conclude.

2. Benefit transfer methodology

2.1. Challenges and prospects

Since the mid 90s, BT has been largely developed and ameliorated, with the main attention progressively switched from the pragmatic aspects of its application to the more methodological ones, e.g. the transfer accuracy, the implications of specific modelling choices for the utility function, the development of more complex prediction methods for meta-analysis and/or for the synthesis of the existing data.

The BT framework is structured upon the so-called "4 S" assumptions that guarantee the provision of consistent benefit estimates at the policy site: 1) separability, 2) specification, 3) sorting and, 4) selection [6]. Separability means that the individual utility is assumed to be separable in its unobserved components which in turn can be captured by the sites and the individuals' characteristics observed by the available data. Specification concerns the correct definition and estimation of the preference function at the study site. It assumes that all the structural components of the utility function are such that the differences in function parameters capture the differences between the study and the policy sites. Sorting assumes that there is not any systematic variation in the unobserved preferences between the sites, such that potential variations are explained by the observable differences in the observed demographic characteristics. Finally, selection assumes that relevant demographic data for the population of the policy site are available/accessible (i.e. data are not affected by selection problems).

The state of the art in BT agrees on the fact that whereas there is no formal test for separability, this is likely to be valid for passive use values and we can use intuition to provide useful guidance in most of the occasions [6]. Neither sorting can be tested but there is agreement on the fact that: 1) it trivially holds whenever the study and policy sites are in the same geographic area and/or referred to the same time and, 2) transfers are related to unexpected events [6]. Specification is potentially testable in terms of the sensitivity of the results to the researcher's modelling choices. However, there is no common agreement neither on the testing procedure to adopt nor on how to evaluate the results consistency. Only meta-analysis constitutes an exception: the widespread idea in BT literature is that such an approach systematically fails in dealing with specification assumption [6]. Finally, selection merely means that either you have data or not.

Within this framework, value transfer is one of the most applied BT approaches by practitioners [36]. It transfers a statistically meaningful WTP value such as the mean or the median WTP from the empirical evidence of a study site to the policy site of interest, as Equation (1) depicts:

$$\widehat{WTP}_g^P = WTP_g^S, \tag{1}$$

where $g = 1, \ldots, G$ indicates the g-th PG and the superscripts P and S indicate the policy and study sites, respectively.

Value transfer exhibits large trade-offs between the simplicity of use and its accuracy. Being based on a point estimate that represents the central tendency of the WTP distribution, when we perform value transfer (either adjusting or not the estimates for the policy site features) we disregard the heterogeneity of the individuals' preferences. If the specification assumption holds and hence the preference function is correctly specified at the study site, by using value transfer we implicitly adjust for differences between the sites [14]. Nevertheless, Ref. [17] assess that a central tendency measure for the WTP distribution is reliable only as far as the study and the policy sites are very similar in terms of both PG features and socio-demographics characteristics. In addition, Ref. [24] highlight that such a situation does not occur in most of the BT applications while Ref. [37], investigating how much the degree of dissimilarity between the sites affects the appropriateness of value transfer, conclude that mean/median value transfer have to be used only when the study and the policy sites are really similar, otherwise function transfer has to be preferred.

Function transfer consists in transferring a function that considers as much relevant information as possible referred to the specificities of the sites, with respect to both the PG features and the individuals' characteristics. The first step of function transfer consists in estimating the WTP model coefficients for the g-th PG at the study site where the elicitation of the individuals' preferences has been carried out, as Equation (2) depicts:

$$\widehat{WTP}_{gj}^S = f\left(\boldsymbol{\beta}_g^S, \boldsymbol{X}_{gj}^S\right), \tag{2}$$

where $\widehat{WTP}^S$ indicates the estimated WTP for the g-th PG at the study site S with respect to the j-th individual. The WTP is obtained as a function of the estimated coefficients beta of the model and the characteristics of both the g-th PG investigated and the j-th individual considered (i.e. the covariates $\boldsymbol{X}$). The second step consists in using the coefficients by applying them for transferring the WTP at the policy site:

$$\widehat{WTP}_{gi}^P = f\left(\boldsymbol{\beta}_g^S, \boldsymbol{X}_{gi}^P\right), \tag{3}$$

where the transferred WTP value is a function of the previously estimated beta coefficients and the covariates $\boldsymbol{X}$ at the policy site. Depending on data availability, the application can concern households and/or target groups instead of individuals as well as it can resort to aggregated measures e.g. the mean household income rather than using covariates at the individual level [38].

Function transfer can be based on different specifications of the functional form [17,37] and more specific modelling can contribute to decrease the average benefit transfer error [18]. The Tobit model, for example, can be applied to exclude non-negative WTP values by means of a threshold, as Equation (4) depicts:

$$\widehat{WTPc}_{gi}^P = \begin{cases} \widehat{WTP}_{gi}^P & \text{if } WTP_{gi}^P \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\widehat{WTPc}^{P}_{gi}$ indicates the values of WTP greater than the threshold, related to the g-th PG and referred to the i-th individual at the policy site.

As Ref. [37] point out, function transfer has two main drawbacks: 1) the function can be easily mis-specified and, 2) it can be easily over-parametrized, meaning that it can fit the study site data but it is inappropriate for the policy site sample. However Ref. [37], conclude that function transfer is generally better able to undertake transfers that consider properly the degree of heterogeneity between the sites, an aspect that plays a crucial role in the transfer results validity. Other authors claim that function transfer basically outperforms value transfer [40,41].

Beyond the empirical evidence that each author offers, Refs. [6,37, 39] stress that the literature in BT is not coherent nor cohesive, being characterized by both a lack of agreement on the best performing transfer approach and a shared BT framework. We consider the above-mentioned issues as unsolved BT challenges, both from the methodological and the practical point of view. In this sense, a cautious approach is suggested also by the debate on the error produced by different BT methods. Usually, BT literature addresses the error of the transfer in terms of the percentage of Benefit Transfer Error (BTE) (we define and discuss BTE in details in section 2.3). Many authors highlight that the percentage of BTE produced by several value transfer applications is equal to 45% (median BTE) and 140% (mean BTE) [13]. These BTE percentages are not lower than the ones produced by several function transfers, where the median percentage is 36% and the mean is 65% but, considering the applications case-by-case, there is conflicting evidence about the approach that has to be elected as the "best performing one" [23], with e.g. extreme cases of BTEs of 900% for function transfer that exist.

The main straightforward point remains, however, that value transfer cannot reproduce the whole WTP heterogeneity at the policy site. It can provide good/bad estimates of the mean or median WTP by transferring a (yet meaningful) representative value for the central tendency of the WTP distribution, but it cannot reproduce it as a whole. In terms of policy planning and adoption, this means that policymakers cannot design the policy to account for target-specific population groups with e.g. peculiar preferences and needs. In turn, this can cause drawbacks in terms of policy efficiency and effectiveness. Function transfer instead can reproduce the heterogeneity but limitedly to the chosen functional form and to the burdens deriving from the assumptions implied by the modelling choice. For example, when it comes to model linearity, function transfer flattens the individuals' preferences heterogeneity, finally distorting the WTP at the policy site and thus leading to similar drawbacks from the policy point of view.

2.2. A novel approach: the Statistical Matching benefit transfer

We propose a BT method that addresses these challenges being able to reproduce the heterogeneity of the individuals' preferences and thus the whole WTP distribution at the policy site. Our approach also lowers the benefit transfer error compared to the traditional BT approaches. This is due to the fact that SMBT can be used to transfer individual values of WTP, meaning that the WTPs elicited for each person interviewed at the study site can be used to predict the WTPs of the individuals at the policy site, preserving the "real", observed information.

We propose to approach benefit transfer through Statistical Matching, as Equation (5) depicts:

$$\widehat{WTP}^{P}_{gi} = WTP^{S}_{gj^*}\ , \tag{5}$$

where the WTP is referred to the g-th PG observed in S for the j^*-th individual that is the most similar one in terms of socio-demographic characteristics to the i-th individual in P. Equation (5) depicts the general SMBT framework. The transfer of the WTP elicited (e.g. by means of contingent valuation) occurs between a study site observation ("the donor") and a policy site observation (the "recipient"), after that these two have been previously paired according to the similarity in some observed characteristics collected by the covariates.

The two fundamental components of the non-parametric micro SM (commonly called "hot deck" methods) are the distance functions and the *techniques*, with the pairing procedure being the core of such a method. The first element (distance function) determines the similarity (or closeness) between the individuals, formally defined as the distance in terms of the observed covariates. The second element (*technique*) is the procedure for the selection, for each individual of the policy site, of the most similar study site individual to be paired with, given the distances previously defined. SM literature developed four *techniques*[42].[1] Three of them are based on distance functions.[2] Among different potential combinations of *techniques* and distances, we chose the Constrained Nearest Neighbor Distance (cnnd) with the Mahalanobis (ms) distance function. This choice is based on the findings and suggestions (about its better performance with respect to the others) in Ref. [44].[3]

To formally describe the SMBT approach, say "matching variables" the covariates $\boldsymbol{X}$ used to assess the similarity between the study and policy site individuals, with $\boldsymbol{X} = \{X1, ., Xl, ., XL\}$, being Xl^S a vector of dimension $(n_S \times 1)$, Xl^P a vector of dimension $(n_P \times 1)$. Superscripts S and P denote the study site and the policy site, respectively. Let be that we have several WTPs per each g-th PG, so that we define the set of variables $\boldsymbol{WTP^S} = \{WTP^S_1, ., WTP^S_g, ., WTP^S_G\}$, being WTP^S_g a vector of dimension $(n_S \times 1)$.

In order to compute the distance among the i-th and j-th observations, we use a distance function. For the sake of simplicity, let be $L = 1$, i.e. we have a univariate (continuous) variable X that is jointly observed between the sites. For the sake of brevity, we refer to Ref. [43] for a generic definition of distance function, limiting us to define here the Mahalanobis distance, as Equation (6) depicts:

$$\Delta^{ms}_{ij} = \sum_{l=1}^{L}\left(X^P_i - X^S_j\right)' \Sigma^{-1}_{X^P X^S}\left(X^P_i - X^S_j\right), \tag{6}$$

where the i-th and j-th observations are distant, with respect to the values of the l-th matching variable X, the amount Δ, that is equal to the sum of the differences between the recipient and donor units, weighted by the covariance matrix Σ of the chosen matching variable.

Given this distance, the cnnd *technique* selects for each recipient the donor that is the closest in terms of the lowest aggregate distance among the distances between the i-th and j-th observations, constraint to the fact that, first, there can be only one donor for each recipient and, second, donors and recipients can be selected only within homogeneous sub-groups defined by the donation classes. These homogeneous groups of donor and recipient units are used to restrict the potential matched units' pairs. Let be the widest potential set of donor and recipient units' pairs defined as $n_S^{n_P}$, where n_S and n_P represent the number of individuals observed in S and P, respectively. Let be X_1 and X_2 two variables that are observed both in S and in P; if a donation class defined upon these two variables holds, the previous set can be restrained to: $(n^S_{X_1})^{n^P_{X_1}} + (n^S_{X_2})^{n^P_{X_2}}$.

The selection of the matching variables is a particularly cumbersome

[1] These *techniques* are the Nearest Neighbor Distance, the Constrained Nearest Neighbor Distance, the Random Distance and the Rank.

[2] There are several distance functions that can be used in the non-parametric SM approach and, among them, the default ones are the Manhattan, the Mahalanobis, the Gower and the Exact ones, as defined by Ref. [43].

[3] Coherently with the results of Ref. [44], the cnnd.ms combination is the best performing one even in the applications presented here. Indeed, we performed the SMBT by applying the Rank *technique* and all the possible combinations among the Nearest Neighbor Distance, Constrained Nearest Neighbor Distance and Random *techniques* and the Manhattan, Mahalanobis and Exact distances, substantiating the findings of [44]. For the sake of brevity, we show and discuss only the results of the best performing combination cnnd.ms.

issue. Counterintuitively, more information added (in terms of a greater number of matching variables selected) to compute the closeness of the respective units does not necessarily increase the power of the matching and does not add robustness [45]. The choice of the matching variables has to be careful and guided by both the specific literature on the research topic and fit-driven criteria. According to the state of the art in SM [46], we choose a multiple linear regression model in order to identify among all the covariates X the best predictors for the variable WTP that we want to transfer from the study to the policy site.

We assume that $XWTP$ (with $XWTP \subseteq X$) is the minimum subset available in S of the variables X that explains the variable WTP in the best way among all the covariates X that are available both in S and P. Being the variables $\boldsymbol{WTP^S}$ continuous, we look at their correlation with the potential predictors and at the adjusted R^2 related to the regression model that explains the individuals' WTP referred to a specific PG in terms of the observed characteristics of the individual [47]. Therefore, accordingly to the predictive power of the variables X on the variable WTP, we select the matching variables to be used, as Equation (7) depicts:

$$WTP^S_g = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_l X_l + \ldots + \beta_L X_L, \tag{7}$$

With $g = 1, 2$, meaning that we define one multiple linear regression model for each WTP (one for each PG). The models can differ with respect to the number of predictors included in Equation (7). These are selected by means of a stepwise procedure in terms of the statistical significance of the coefficients, the R^2, the adjusted R^2 and the F-statistics. Among all the predictors that the models have in common, we choose the most relevant ones looking at their statistical significance and at the hints from the economic theory. Following Ref. [46], we also select among the covariates X some categorical variables that are relevant according to the economic theory in explaining the dependent variables $\boldsymbol{WTP^S}$, using them to build the donation classes.

2.3. Benefit transfer methods comparison

In order to assess the validity of the different BT approaches, we consider the convergent validity criterion that is largely used in BT literature [6,22–25]. Since there is not a consensus on which approach to convergent validity has to be adopted [6], we choose to quantify the benefit transfer error as in Ref. [23] which is also computed by the most recent BT literature [16,17,36,37,48,49].

The BTE is computed as Equation (8) depicts:

$$\left|\%BTE\right| = \left[\left(\widehat{WTP^P_{gi}} \big/ WTP^P_{gi}\right) - 1\right] \times 100, \tag{8}$$

where the numerator is the WTP transferred to the policy site from the study site and the denominator is the WTP elicited at the policy site, both referred to the g-th PG in relation to the i-th individual. Then, we consider a mean measure of BTE.

In addition, considering that the BTE is referred to an average (relative) error [6] but we propose a BT approach that transfers individual WTP values, we include in the comparison procedure two other tools for the assessment of the BT validity: 1) the validation strategy proposed by Ref. [44] and, 2) the Hellinger index [50]. This choice is due to the fact that the BTE measures the error produced, on average, by BT. Whereas its minimization is a more than suitable goal to aim at, the main focus of the present paper is to investigate the ability of BT to reproduce the WTP heterogeneity at the policy site and to reflect the actual WTP distribution as a whole. Then, we are interested in investigating (and quantifying) the dissimilarity between the WTP distributions: the actual one at the policy site and the one reconstructed by transferring the individual information from the study site. An average error minimization criterion cannot serve this purpose properly. Therefore, the assessment of the transfer validity is tackled by considering the current BTE estimation procedure as benchmark but also the two tools derived from the data integration context.

Finally, being that by SMBT we transfer individual WTP values and thus the whole WTP distribution, but also considering that we operate in a non-parametric framework, we calculate the empirical bounds of the transferred distributions. The empirical bounds give insights into the transfer uncertainty and they are determined by drawing 1000 random samples of the transferred distributions (i.e. of the individual WTP values) by randomly dropping out every time the 25% of the study site/donor sample (leave-k-out strategy).

3. Application and results

3.1. Case study

We transfer the WTPs elicited for soil erosion and carbon sequestration, two relevant PGs provided by agriculture and forestry in the hilly and mountain areas of the Italian region of Emilia-Romagna. An online survey was carried out by a professional agency during December 2016 targeting the regional population (based on residential quotas for age classes, gender and provinces), after having pre-tested it on a pilot.

The contingent valuation method was used to elicit the WTP. Respondents were asked to answer to a single open-ended question for each PG, after having answered to several questions related to the areas and the goods under analysis. Hence, the elicitation protocol was based on the contextualization of each PG first, in terms of its general definition, second, in terms of the effect produced in the areas of interest and, third, in terms of their measurement (also providing information on the current level of provision). Preliminary work with agri-environmental authorities and regional stakeholders was dedicated to the appropriate and reasonable definition of each PG. The respondents were asked about the WTP for reaching, annually, the optimal level of soil erosion reduction and the optimal capacity of carbon sequestration. Moreover, respondents had to express a preference with respect to the payment mechanism to be employed for corresponding the WTP. Appendix A depicts the section of the questionnaire (translated in English from Italian) dedicated to the contingent valuation exercise. The questionnaire included also 5-points Likert-scale attitudinal questions related to the respondents' relation with the hilly and mountain areas of Emilia-Romagna, their perceptions on the relevance of the PGs under analysis and the main issues that they think are related to them. Respondents were also asked about their socio-demographic characteristics and information on their household expenditures, e.g. the annual payment for a food basket, the current payment for the Land Reclamation Authority (LRA), etc.

We identify two different applications of BT by ex-post dividing the whole sample of 1007 final valid respondents in two sites. Fig. 1 depicts the areas considered. In Application 1, the policy site is the province of Bologna while the study site is represented by the other provinces of Emilia-Romagna. In Application 2, the policy site is the province of Ferrara while the study site is represented by the other provinces of Emilia-Romagna. Table A1 and Table A2 show the summary statistics for the respondents' samples considered in the two applications.

Among the covariates observed for the two samples (listed in Table A3), we select the ones to include in the linear and censored Tobit models that are applied for function transfer. This choice is made according to which covariate the theory suggests to be (at least potentially) the determinants of the WTP [13,18,37]. The final models presented in sections 3.2.2 and 3.2.3 result from both theory-driven and fit-driven rationales.

3.2. WTP transfer

3.2.1. Value transfer

In both Application 1 and Application 2, we apply value transfer transferring both the mean and the median WTP elicited in the two respective study sites. We stress that in both the applications no

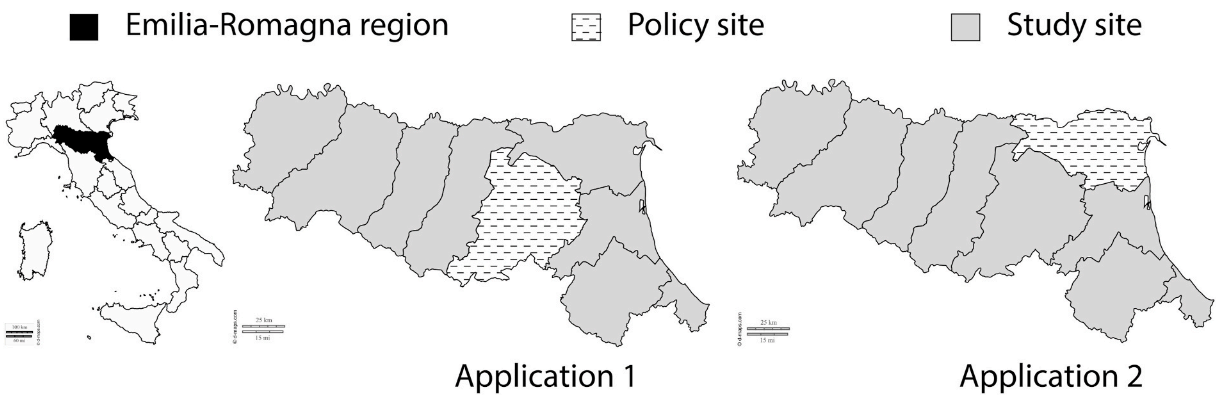

Fig. 1. Map of Emilia-Romagna region (Italy) with the areas under analysis.

statistically significant differences hold among the true mean or median WTPs at the policy site and the ones transferred from the respective study site.

3.2.2. Function transfer: linear model

The (final) multiple linear regression model selected for the function transfer of the WTP for soil erosion (in both the applications) is depicted by Equation (B.1) while Table B1 shows the coefficients of the models resulting from the two applications. The (final) multiple linear regression model selected for the function transfer of the WTP for carbon sequestration (in both the applications) is depicted by Equation (B.2) while Table B2 shows the coefficients of the models.

In Application 1, the models for both the PGs have a positive intercept (not statistically significant) while the most significant coefficients are related to the presence of a farmer in the family (*household_farmer)*, the amount of payment for the LRA (*lra_pay)* and the preferred payment modality (*how_wtp_pay*). Being aware of the "soil erosion issue" and having personal connections with the hilly and mountain areas are highly correlated to the WTP, as expected. Both the beliefs that land abandonment positively affects soil erosion (*aband_eros_p*, i.e. that the abandonment of agricultural lands contributes to increase soil erosion) and that land abandonment negatively affects it (*aband_eros_n*, i.e. that it contributes to mitigate soil erosion) lead to an increase of the WTP for the PG. Finally, respondents that are involved in some peculiar activities in the hilly and mountain areas of the region, such as fishing, hunting and others recreational activities, show higher WTP for soil erosion. In Application 2, the models for both the PGs have very similar features compared to the models in Application 1 (sign, magnitude, statistical significance of the coefficients as well as the reasonably acceptable pseudo-R^2).

3.2.3. Function transfer: tobit model

The (final) censored Tobit regression model selected for the function transfer of the WTP for soil erosion (in both the applications) is depicted by Equation (B.3) while Table B3 shows the coefficients of the models resulting from the two applications. The (final) censored Tobit regression model selected for the function transfer of the WTP for carbon sequestration (in both the applications) is depicted by Equation (B.4) while Table B4 shows the coefficients of the models.

In Application 1, the models for both the PGs show positive intercepts that are highly statistically significant. The presence of a farmer in the household is still highly significant with respondents that both pay more for the LRA and have a higher income, showing a higher WTP for soil erosion and carbon sequestration. In the model for soil erosion, older respondents tend to pay less for the PG. In Application 2, the models for both the PGs show positive intercepts that are highly statistically significant, with the presence of a farmer in the household that is highly significant as well as the amount of payment for the LRA and the income level (i.e. respondents with higher income and that pay more for LRA tend to have higher WTP for soil erosion and carbon sequestration).

3.2.4. Statistical Matching benefit transfer

The selection procedure of the matching variables to be used by SMBT is guided by the statistical significance of the determinants of the dependent variable (i.e. the WTP) of the multiple linear regression models on soil erosion and carbon sequestration, respectively. The two best-fitted linear models in terms of the coefficients and the pseudo-R^2 are chosen and, among the finally included predictors, we select two continuous variables (*lra_pay* and *household_income*) that are both relevant in the estimation of the PGs WTP as well as they are easily at disposal for the policymakers in real-life circumstances. The donation classes are defined by means of three dummy variables, namely: *mount_house*, *gender* and *employ*. They are used, as prescribed by the literature in SM [42], to define some homogeneous groups (with respect to the respondents' residence in hilly and mountain areas, the working status and the gender) for the transfer.

3.3. Comparison among the benefit transfer approaches

Fig. 2 (A) shows the distribution of the WTP for soil erosion while Fig. 2 (B) shows the distribution of the WTP for carbon sequestration as they are reproduced at the policy site by means of the different BT methods in Application 1.

Considering the WTP distribution as a whole, for both the PGs under analysis this results to be flattened around the mean/median WTP. This happens in all the approaches but the SMBT. In particular, function transfer with linear model specification implies two main potentially undesired features. First, we transfer also negative WTP estimates (i.e. individual WTPs for soil erosion and carbon sequestration that are below 0€) whereas these values are not originally elicited at the study site. Second, the WTP values that are both close to 0€ (representing potential protesters and/or people that express an indifference or a really low interest with respect to the two PGs) and close to 300€ are largely underestimated. If the censored model allows us to avoid the first drawback, it does not reproduce the higher WTP values (i.e. the right tails of the WTP distributions). SMBT outperforms the other approaches in reproducing the original multimodal WTP distribution at the policy site, thus properly characterizing the WTP for the three groups of individuals that, ideally: 1) do not pay or pay only a small amount of money, 2) do pay an average amount of money and, 3) do pay a lot. By applying SMBT we do not overestimate the WTP values close to the average WTP, as we do instead by means of function transfer (with both linear and Tobit model specifications). Therefore, the true heterogeneity of the individuals' preferences at the policy site is not flattened while the really low and really high values of WTP are robustly estimated.

Fig. 3 (A) and Fig. 3 (B) depict the quartiles of the WTP distributions for soil erosion and carbon sequestration, respectively, as they are reproduced at the policy site by means of the different BT methods in

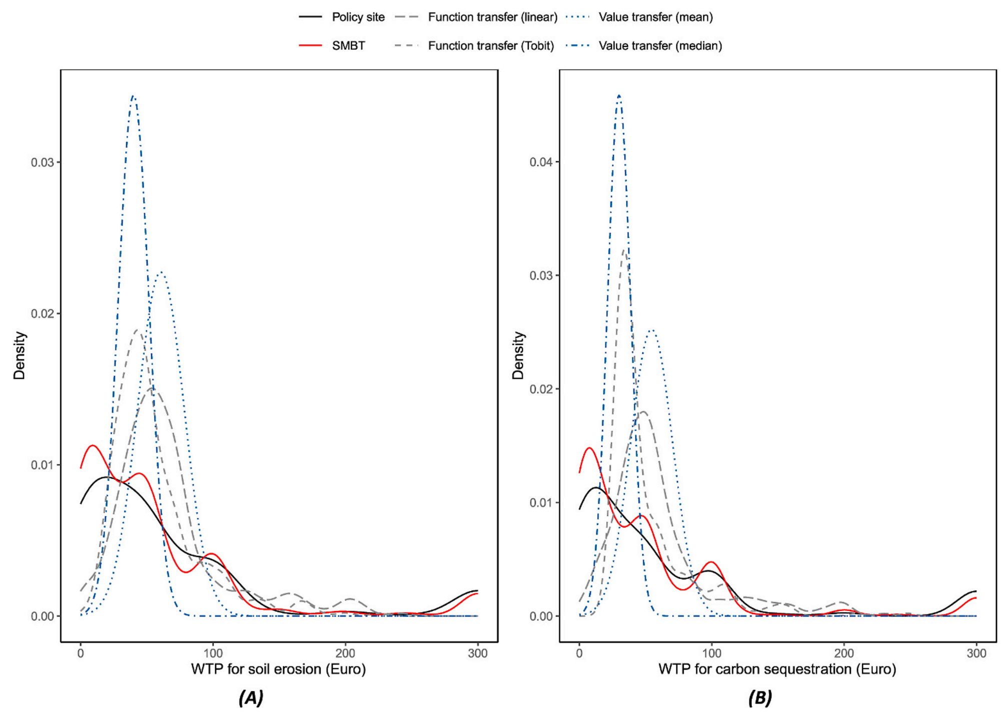

Fig. 2. Different benefit transfer methods compared in Application 1. (A) WTP for soil erosion. (B) WTP for carbon sequestration.

Application 1.

The boxplots depict the median WTP (bold vertical line within the box) as well as the first and third quartiles (lower and upper lines delimiting the box). The size of the box indicates the dispersion in the data at hand while the whiskers that extend from the box indicate the variability outside the first and third quartiles. Boxplots are non-parametric graphical devices particularly suitable to display the variation of the sample without making any assumptions of the underlying statistical distribution.

Considering the mean/median WTP, by means of value transfer, the transferred WTP for soil erosion is 60.58€ (mean) or 40€ (median) while the transferred WTP for carbon sequestration is 54.57€ (mean) or 30€ (median). Function transfer with linear model specification allows to transfer individual WTP values from which the resulting mean WTP is 63.68€ (55.47€ the median) for soil erosion while for carbon sequestration it is 57.69€ (49.42€ the median). Individual WTP values transferred by function transfer with Tobit model specification bring to a mean WTP of 57.84€ (47.27€ the median) for soil erosion while for carbon sequestration it is 52.80€ (38.36€ the median). With SMBT the resulting mean WTP for soil erosion is 55.40€ (40€ the median) while it

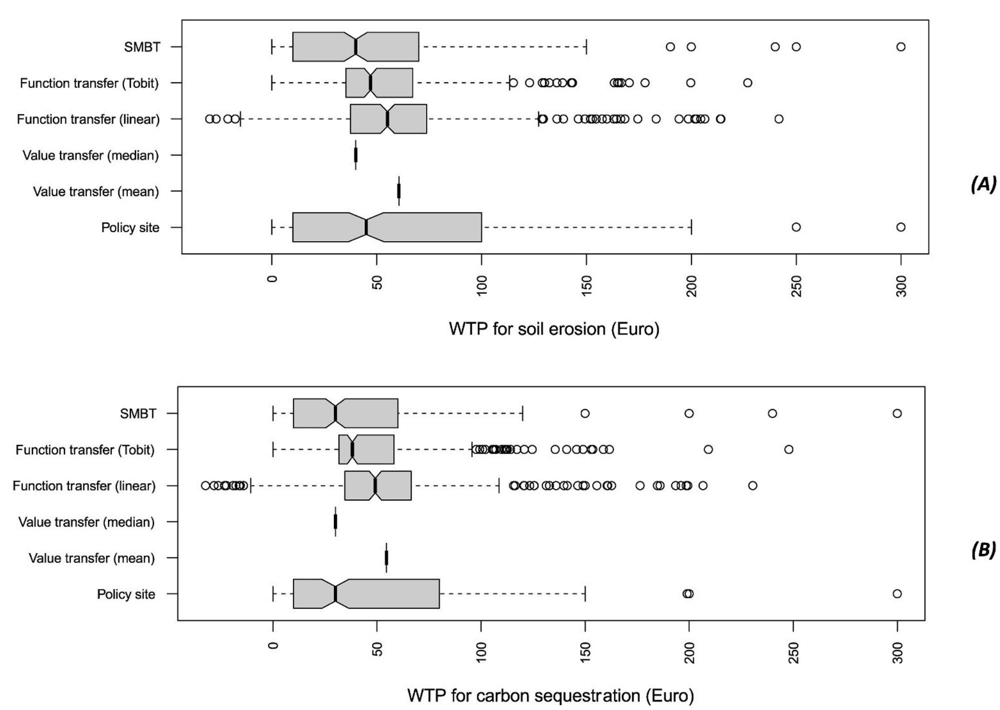

Fig. 3. Boxplots of WTP distribution reproduced by different benefit transfer methods in Application 1. (A) WTP for soil erosion. (B) WTP for carbon sequestration.

is 49.73€ (30€ the median) for carbon sequestration. These average measures of WTP for both the PGs are not statistically significant different (at 10% for Tobit model specification, 5% for all the others) from the means originally elicited at the policy site (Bologna province). In this latter, the originally elicited mean WTP for soil erosion is 64.97€ (45€ the median) while for carbon sequestration it is 61.33€ (30€ the median), as Table A1 depicts.

Fig. 3(A) and (B) show that: 1) mean value transfer as well as function transfer (both linear and Tobit model specifications) flatten the reproduced WTP around the mean (while median value transfer flattens it around the second quartile); 2) the dispersion of the values between the first and the third quartiles is shrunk and, in the worst cases, it can be nullified (mean/median value transfer); 3) the variability of the WTP towards the higher values (right whiskers) is mis-reproduced, if not by means of outliers by applying function transfers. Moreover, linear function transfer produces also negative estimates of WTP that are not elicited at the policy site, originally. SMBT does not produce such distortions. It outperforms the other methods in reproducing the central tendency of the WTP distribution as well as the variability of the lower and higher values (i.e. the left and right whiskers departing from the box). SMBT is also the only method that reproduces the upper extreme outliers. These considerations are valid with respect to both the WTP for soil erosion and carbon sequestration.

Fig. 4 (A) shows the distribution of the WTP for soil erosion while Fig. 4 (B) shows the distribution of the WTP for carbon sequestration as they are reproduced at the policy site by means of the different BT methods in Application 2.

As in Application 1, also in Application 2, when we consider the WTP distribution as a whole, for both the PGs, by applying BT methods others than SMBT, we fail to reproduce the actual distributions of the policy site and all the drawbacks from Application 1 do hold. Moreover, in Application 2 these drawbacks are even intensified due to the increased degree of dissimilarity between the study site and the policy site of interest.

Fig. 5 (A) and Fig. 5 (B) depict the quartiles of the WTP distributions for soil erosion and carbon sequestration, respectively, as they are reproduced at the policy site by means of the different BT methods in Application 2.

Considering the mean/median WTP, by means of value transfer, the transferred WTP for soil erosion is 62.68€ (mean) or 40€ (median) while the transferred WTP for carbon sequestration is 57.24€ (mean) or 30€ (median). Function transfer with linear model specification allows to transfer individual WTP values from which the resulting mean WTP is 52.21€ (57.07€ the median) for soil erosion while for carbon sequestration it is 49.43€ (51.46€ the median). Individual WTP values transferred by function transfer with Tobit model specification bring to a mean WTP of 54.84€ (52.50€ the median) for soil erosion while for carbon sequestration it is 49.81€ (40.52€ the median). SMBT produces, by means of the individual values transferred, a mean WTP for soil erosion of 67.38€ (40€ the median) while it is 56.79€ (30€ the median) for carbon sequestration. These average measures of WTP for both the PGs are not statistically significant different (at 5%) from the means originally elicited at the policy site (Ferrara province). Here, the originally elicited mean WTP for soil erosion is 50.26€ (30€ the median) while the originally elicited mean WTP for carbon sequestration is 46.54€ (22€ the median), as Table A2 depicts.

Also in Application 2 (that is characterized by an increased degree of dissimilarity between the study site and the policy site of interest), Fig. 5 (A) and (B) show that: 1) mean value transfer as well as function transfer (both linear and Tobit model specifications) flatten the reproduced WTP around the mean (while median value transfer flattens it around the second quartile); 2) function transfer (both linear and Tobit model specifications) distorts the dispersion of the values between the first and the third quartile, not only by shrinking it but also by inverting the interquartile spread of the distribution (i.e. they tend to mis-reproduce the box asymmetry). In the worst cases the dispersion can be nullified (mean/median value transfer). 3) The variability of the WTP towards the lower values is mis-reproduced (left whiskers). Moreover, linear function transfer produces negative estimates of WTP that are not elicited at the policy site, originally. As in Application 1, also in Application 2 SMBT does not produce such distortions. It outperforms the other methods in reproducing the central tendency of the WTP distribution, the variability of the lower values, the interquartile dispersion as well as the reproduction of the upper extreme outliers. These considerations are valid with respect to both the WTP for soil erosion and carbon sequestration.

We consider now the convergent validity criterion that is largely used in BT, the Benefit Transfer Error and, successively, the Hellinger index. The former allows to quantify the average error behind the transfer procedure, the latter offers a measure of dissimilarity among the distributions. Table 1 shows the BTE for the different BT approaches in both the applications. In Application 1, the value transfer based on the median WTP outperforms the other approaches with respect to both the PGs. In Application 2, instead, it is the SMBT approach (that is also the second-best performing BT method in Application 1) that outperforms all the others.

Table 2 shows the measures of the Hellinger index for the different BT approaches. We consider this index in order to quantify the

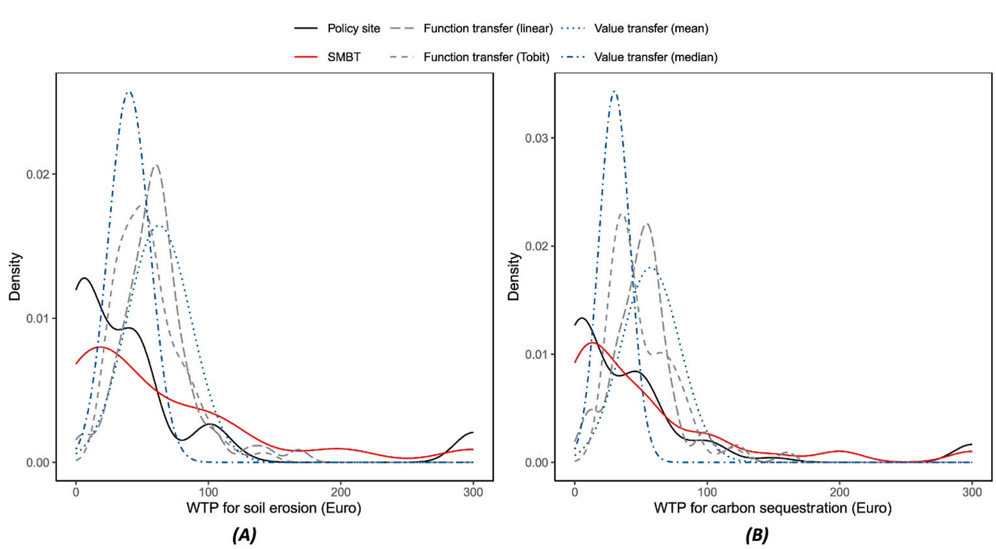

Fig. 4. Different benefit transfer methods compared in Application 2. (A) WTP for soil erosion. (B) WTP for carbon sequestration.

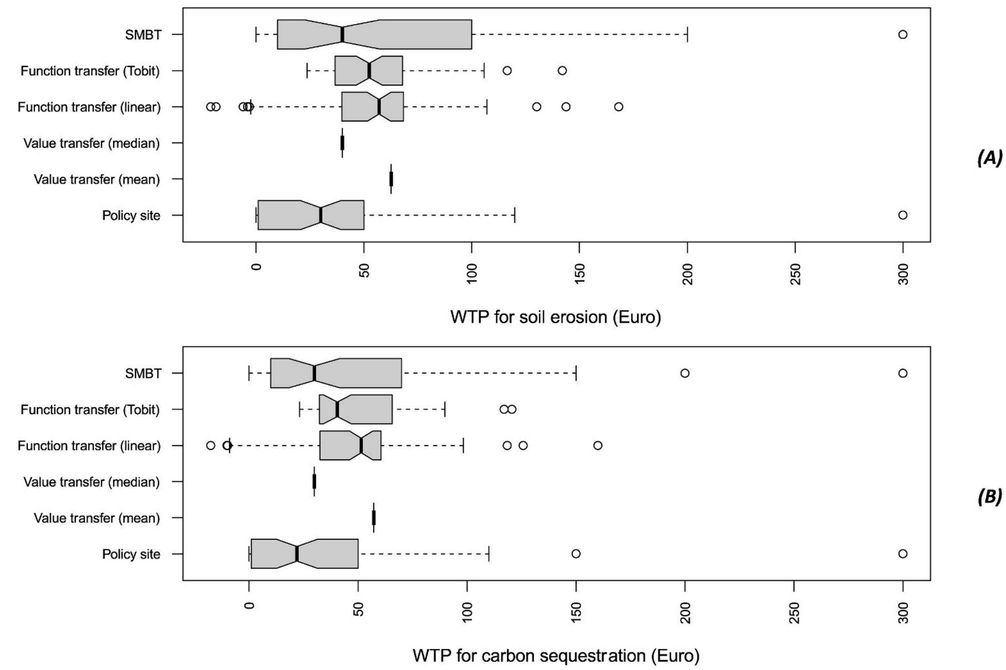

Fig. 5. Boxplots of WTP distribution reproduced by different benefit transfer methods in Application 2. (A) WTP for soil erosion. (B) WTP for carbon sequestration.

Table 1
Comparison among the benefit transfer errors generated by different benefit transfer methods.

Benefit transfer errors				
Benefit transfer method	Application 1 (policy site: Bologna province; study site: other provinces)		Application 2 (policy site: Ferrara province; study site: other provinces)	
	Soil erosion	Carbon sequestration	Soil erosion	Carbon sequestration
Value (mean)	209.356%	250.952%	394.237%	598.397%
Value (median)	108.473%	98.695%	223.366%	285.853%
Function (linear)	183.676%	201.318%	396.526%	526.997%
Function (Tobit)	152.126%	192.117%	349.326%	459.271%
SMBT (cnnd. ms)	115.122%	154.807%	167.688%	273.358%

SMBT: Statistical Matching benefit transfer. Cnnd.ms: Constrained Nearest Neighbor Distance technique with Mahalanobis distance function combination.

Table 2
Comparison among Hellinger index measures generated by different benefit transfer methods.

Hellinger distance index				
Benefit transfer method	Application 1 (policy site: Bologna province; study site: other provinces)		Application 2 (policy site: Ferrara province; study site: other provinces)	
	Soil erosion	Carbon sequestration	Soil erosion	Carbon sequestration
Value (mean)	0.21249	0.24724	0.17959	0.15049
Value (median)	0.26203	0.30586	0.15513	0.20167
Function (linear)	0.08999	0.10925	0.14253	0.09185
Function (Tobit)	0.09604	0.17528	0.14152	0.13802
SMBT (cnnd. ms)	0.00743	0.01137	0.04234	0.02592

SMBT: Statistical Matching benefit transfer. Cnnd.ms: Constrained Nearest Neighbor Distance technique with Mahalanobis distance function combination.

dissimilarity between the distributions of the true WTP at the policy site and the one transferred from the study site. A 5% threshold is conventionally considered to be the maximum acceptable amount of dissimilarity [50]. Value transfer (both using the mean and the median WTP) produces the largest degree of dissimilarity in relation to both soil erosion and carbon sequestration in both the applications.

In Application 1, the dissimilarity between the two WTPs when we transfer the mean is equal to 26% and 30% for soil erosion and carbon sequestration, respectively. It is equal to 21% and 24% when we transfer the median WTP for soil erosion and carbon sequestration, respectively. Function transfer produces almost 9% (soil erosion) and 11% (carbon sequestration) of the dissimilarity when the transfer is performed with the linear model. The censored (Tobit) model produces 9.6% (soil erosion) and 17% (carbon sequestration) of the dissimilarity. Greatly superior is the SMBT that produces only 0.74% of the dissimilarity for the transfer of the WTP for soil erosion and 1.13% of the dissimilarity for the transfer of the WTP for carbon sequestration.

In Application 2, the dissimilarity between the two WTPs when we transfer the mean is equal to 18% and 15% for soil erosion and carbon sequestration, respectively. It is equal to 15% and 20% for soil erosion and carbon sequestration, respectively, when we transfer the median. Function transfer produces 14% (soil erosion) and 9% (carbon sequestration) of the dissimilarity when BT is performed with the linear model. The censored (Tobit) model produces 14% (soil erosion) and 13% (carbon sequestration) of the dissimilarity. Again, the SMBT is greatly superior, producing 4.2% of the dissimilarity for soil erosion and 2.5% of the dissimilarity for carbon sequestration. Note that in Application 2 the SMBT is the only method that produces an acceptable amount of dissimilarity (under the threshold of 5%).

Appendix C displays the results of the leave-*k*-out strategy with *k* equals to ¼ of the original study site/donor sample related not only to the proposed SMBT (Fig. C3) but also to function transfer with both the linear model specification (Fig. C1) and the censored Tobit model specification (Fig. C2), for both the PGs under analysis in both Application 1 and Application 2. Looking at Fig. C3, the transfer uncertainty behind the SMBT is well-contained. The empirical bounds of the WTP distributions transferred by SMBT both in Application 1 and Application

2 and the comparison between them and the ones in Fig. C1 and Fig. C2 spread some light on the fact that: 1) potentially, SMBT can perform even better in representing the WTP distribution tails than the mere results discussed previously and, 2) SMBT performs well in containing the uncertainty behind the transfer when the degree of similarity between the sites decreases (as it happens moving from Application 1 to Application 2).

4. Discussion

First of all, we contextualize the estimated PGs WTPs with respect to the state of the art.

On the one hand, considering soil erosion, to the best of our knowledge few other works provide insights on such a peculiar PG. Indeed, the most part of the existing papers focus on aggregate policy instruments like e.g. soil and water conservation measures [51]. These measures consider altogether different PGs like watershed preservation, reduction of soil erosion, increasing/preservation of soil fertility, maintaining of the irrigation water availability, etc. Due to the fact that a unique payment is elicited for all of them, comparisons among WTP values are difficult to make. In Ref. [52], a choice experiment in two watersheds of Southern Spain is carreid out, eliciting in three different scenarios a mean WTP for soil erosion that ranges from 26.23€ up to 63.61€ (per person/per year). In Ref. [53], by a contingent valuation exercise in Andalusia, a mean WTP for soil erosion in different scenarios that ranges from 11€ up to 53€ (per individual/per year) is estimated.

On the other hand, carbon sequestration has been investigated especially in the context of wide forests/parks (e.g. the Amazon) with respect to the conservation fees that are charged to the incoming tourists with the purpose of guaranteeing the forest capacity of carbon dioxide sequestration. In Ref. [54], by applying a contingent valuation in a Bolivian reserve a mean WTP (per person/per visit) of 36.73US$ (around 34€) is estimated. In Ref. [55], they estimate a mean WTP (per person/per visit) of 15.50US$ (around 14€) carrying out a choice experiment in a big reserve in Peru. In the agro-forestry system of the American households they elicit by choice experiment a mean WTP for carbon sequestration (per household/per year) of 58.05US$ and 62.72US$ in two different scenarios (equal to 53.30€ and 57.56€, respectively) [56].

Taking into account these results, we consider our WTP estimates for the two PGs reliable and valid, basically in line with the existing literature on stated preference methods related to soil erosion and carbon sequestration. This, both considering the mean WTP values originally elicited at the study and policy sites and the mean WTPs resulting from the individual values transferred by the different BT approaches.

In relation to the BT framework, the separability and sorting assumptions can be intuitively and factually considered to hold. The former holds due to the fact that our application considers mainly passive use values. Moreover, regional stakeholders and practitioners who have been consulted after the elicitation exercise to discuss our results expressed satisfaction and agreement on the obtained estimates, considered reliable and accurate for the area under analysis. The latter assumption holds due to the fact that the transfer is performed within the same geographic area (an Italian region), at the same reference time and in relation to unexpected events.

Considering the methodological aspects, SMBT performs the best in terms of BTE in Application 2 and it is second among the applied approaches in Application 1. It is the best performing approach considering the Hellinger index in both the applications. These results hints at the potentiality of the method and help in discussing the drawbacks of value transfer (both with the mean and the median WTP) but also those of function transfer (both linear and Tobit model specifications). The real data from the Emilia-Romagna case study show that, in Application 1, there is a larger degree of similarity between the study site and the policy site with respect to Application 2. This degree of similarity concerns the PGs features and the characteristics of the individuals as well as the originally elicited WTP distributions. The study and policy sites are instead much more different in Application 2 than in Application 1. These sites differences are depicted by the descriptive statistics of Table A1 and Table A2. In addition, focusing on the original WTP distributions for both the PGs (as they have been originally elicited at the policy site) they are much more diverse in Application 2 than Application 1. The variation coefficient of the true WTP is 1.211 (soil erosion) and 1.322 (carbon sequestration) in Application 1 while it is 1.534 (soil erosion) and 1.542 (carbon sequestration) in Application 2. The originally elicited WTPs present also different degrees of asymmetry in their distributions. In Application 1 the skewness is 1.997 (soil erosion) and 2.068 (carbon sequestration) while in Application 2 it is 2.460 (soil erosion) and 2.618 (carbon sequestration). The kurtosis is 6.259 (soil erosion) and 6.462 (carbon sequestration) in Application 1 while it is 8.310 (soil erosion) and 9.569 (carbon sequestration) in Application 2. These indexes suggest that the originally elicited WTPs at the policy sites are very dissimilar between the two applications, with Application 2 that is characterized by a higher degree of diversity with respect to the PGs WTPs.

These results are strictly related to the findings and conclusions in Ref. [37] that suggest that whenever the transfers involve similar provision changes, related to similar PGs, analyzed in similar sites, value transfer should be preferred to other approaches. If any of these similarities ceases to hold, the validity and robustness of value transfer have to be questioned. The greater the similarity between the sites, the better the validity of value transfer, the lower the similarity between the sites, the better the validity of function transfer. However, while these considerations are expected to be valid *a priori*, there is no guarantee that, holding a high degree of dissimilarity between the sites, function transfer (even by means of complex functions) can perform well (e.g. in terms of BTE) when applied to the policy site sample [37]. This argument is supported also by Ref. [17] that is in favor of general functional forms (which can lead to lower transfer errors) in spite of ad hoc explanatory variables and site-specific characteristics. But also the degree of heterogeneity between the sites is a very relevant issue that concurs in determining the success of the transfer in terms of results validity and accuracy [17].

We argue that when the originally elicited WTP distribution is very heterogeneous, it presents a complex multimodal form and it is peculiarly tailed (as it is in Application 2) clearly, we cannot reproduce the heterogeneity of the preferences at the policy site by means of the mean (or median) value transfer. But we also argue that neither function transfer is able to properly accomplish this task. Its main ability indeed is to transfer values that are close to the average WTP. Yet meaningful, such a value cannot give insights on the WTP values that are e.g. in the tails of the distribution. By means of function transfer we then risk to smear the WTP values around the central tendency measure, finally tending to misrepresent the WTP distribution. Even more severe drawbacks potentially arise when a high degree of diversity exists. This is very important also because typically the WTP has a multimodal distribution that is often mis-reproduced by linear and/or censored models while the SMBT can properly adapt to its features.

Fig. 4 (A) and 4 (B), if compared to Fig. 2 (A) and 2 (B), help to clarify this point. The former figures show how the originally elicited WTP distributions, that are less smoothed and much more tailed than the originally elicited WTPs depicted by Fig. 2 (A) and 2 (B), cannot be properly reproduced by the measures of central tendency or the values clustered around the average WTP. Instead, SMBT allows to catch the heterogeneity of the preferences and to reflect the actual WTP distributions. This is even clearer by looking at the Hellinger index: SMBT outperforms all the other approaches in reproducing as much as possible the current, "true" WTP distributions.

Finally, another important result achieved by SMBT is the lowering of the transfer errors. While the improvement of BT accuracy is considered still a pending issue by BT literature [18], in terms of the classic BTE, SMBT produces the second-lowest transfer error in

Application 1 while it largely outperforms all the other approaches in Application 2. These findings are in line with the conclusions of Refs. [14,37] about the relevance in BT of the similarity degree between the study and the policy sites. Figure C3 depicts the empirical bounds for the WTP distributions transferred by SMBT. This latter is particularly able to reproduce the WTP distributions at the policy site by containing the uncertainty levels around the estimated individual values, potentially performing even better than the mere results discussed in section 3.3.

5. Conclusions

In this paper we propose a novel approach to BT that is based on the non-parametric micro SM. Originally envisioned for data integration purposes, this method can be used in benefit transfer in order to transfer individual WTP values to the policy site, as they are observed for the most similar individual present in the study site, only by means of the real, observed information. This allows to account for the heterogeneity of the WTP at the policy site and preserve the whole WTP distribution. Also, the transfer errors are lowered or contained.

While the application of function transfer requires additional effort and evidence of its performance is mixed, particularly when compared with value transfer [17,25,37], SMBT exhibits a more positive balance between the complexity of data used and the improvement in the transfer validity in terms of both the reduction of the transfer errors and the dissimilarity between the WTP distributions. This exploratory application shows that SMBT works fine with a common and easily available/accessible set of covariates such as the average annual expenditure of the households, their average monthly income, the average annual expenditure for the LRA. These are easy-to-find information that can be collected e.g. by means of administrative registers.

There are three main policy-related sets of considerations arising from this study: a) the potential for improved policies using SMBT; b) the implications for data policies related to PGs provision measures; c) the accounting of uncertainty in public policy.

Concerning point a), assuming the problem is to design a policy instrument based on a flat rate payment or taxes, or even regulation, SMBT, through a more accurate estimate of the average WTP will allow to better assess the decision to implement or not the measure. This holds in particular in cases of major differences between the policy and the study sites and when the true WTP has a multimodal distribution and/or when it is particularly tailed. SMBT can however better serve policy design if it is used to construct a proper demand function by ordering the WTPs and, if needed, by weighting the WTPs based on the composition of the population. This allows to better elaborate on the optimal level of PG provision and related policy targets. Moreover, as it allows to better identify different beneficiary groups, potentially expressed by the multimodal distribution of the WTP, SMBT exploits its potential in support of policies more oriented to answer to the differentiated needs of the beneficiaries. This is the case, for example, of Payments for Ecosystem Services, of narrowly targeted interventions (if beneficiaries are highly clustered in space or by typology), of differentiated measures designed according to the benefits perceived or even market-based approaches (including those linked to value chain exploitation of PGs, based on individual consumers WTP and highly dependent on marketing strategy design). This list includes indeed most of the new instruments proposed for the provision of PGs by the current CAP reform debate, in order to improve efficiency and value from money from public expenditure. The potential for policy design prescription in this case may be boosted by means of a further treatment of the transferred individual WTP through cluster analysis or similar classification means. Clustering associated to the characteristics of the groups can also be the basis for designing collective multi-actor approaches, in particular identifying rules about the most preferable groups of actors to be involved in the collective schemes. Grouping or individual marginal benefit can also support a better use of result-based instruments, through a more accurate definition of both targets and remuneration of the results (in other words, identifying more accurately what is the marginal value of the results achieved).

Concerning point b), we recognize that the current level of data availability is far from allowing an easy large-scale exploitation of this approach. However, our results encourage higher investment in data collection and storage. It also corroborates current trends in terms of data policy, in particular concerning open data, data transparency, linkages and interoperability across available datasets. A better data availability in the future will certainly allow a wider exploitation and a better implementation of the method.

Finally, concerning point c), we need to acknowledge that the use of value transfer in policy making is a field in which strong assumptions may yield the illusion of certitude even with poor data and methods. This may affect the quality of decisions dramatically. Especially in a "Post-truth World" [57] and in a field characterized by a high instability of preferences, the SMBT approach can help in improving the quality of policy support related to public goods by better communicating uncertainty about WTP and feed this information into measure design. This also goes in the direction of supporting a cultural shift towards a higher recognition of the (constructive) role of uncertainty in policy support [57].

Further developments of the work should explore the transfer of aggregate values of WTP referred e.g. to groups of individuals. Further developments may also concern the elicitation method adopted to estimate the WTP extending the method to choice experiments. For example, it could be worthy to explore how to adapt SMBT to consider the transfer of the whole set of answers to the choice cards that has been proposed to the study site by different respondents. Finally, a somehow straightforward and complementary extension of this work is the BT applied to the supply side, in order to better reproduce the heterogeneity of farmers' compliance costs and their distribution [58]. This may help estimate more precise supply functions and improve cost-targeting in order to increase effectiveness and efficiency of public expenditure.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Contingent valuation exercise and features of Application 1 and Application 2

In the section ***Part 2 – Economic valuation*** of the questionnaire, the respondent is asked to contribute to evaluate some potential scenarios of improvement of the production of two PGs in the hilly and mountain areas of Emilia-Romagna.

Soil erosion

First, the PG is defined in its most general meaning, as follows:

"Soil erosion, in the hilly and mountain areas, is conceived as the phenomenon of removal of the foundational material of the local area"

Second, the PG is described in terms of its expected effects, as follows:
"The effects of soil erosion in the hilly and mountain areas are the reduction of agricultural land fertility and the increasing of the slopes instability. Eroded soil is then transported down to the valley where it contributes to decrease the efficiency and the flow rate of rivers and canals with the consequent increasing of floods risk"

Third, it is defined how the PG is measured, as follows:
"Soil erosion is measured in terms of the quantity of soil that is annually lost in the hilly and mountain areas that are defined as stable, i.e. the areas that are not interested by landslides"

Fourth, information on the actual level of provision of the PG in the area of interest is given:
"In hilly and mountain areas of Emilia-Romagna, every year, there is a loss of around 14 million tons of soil, corresponding to around 21.4 tons per hectare of stable land in the hilly and mountain areas"

Fifth, the open-ended question to elicit the WTP is asked, as follows:
"How much is the maximum annual amount that your family is willingness to pay in order to reach the optimal level of reduction of soil erosion in the hilly and mountain areas?"

Sixth, if the respondent is willing to pay, he/she is asked to express with respect to the previous payment his preference in terms of payment modality, choosing among:

- *Generic private body*
- *Public authority* ad hoc
- *Increase of current taxation*

Carbon sequestration

First, the PG is defined in its most general meaning, as follows:
"Carbon sequestration, in the hilly and mountain areas, is conceived as the amount of carbon dioxide (CO2) sequestered in terms of wood and vegetation by the forestry"

Second, the PG is described in terms of its expected effects, as follows:
"The carbon sequestered from the atmosphere contributes to mitigate the climate change"

Third, it is defined how the PG is measured, as follows:
"Carbon sequestration is measured in terms of the quantity of carbon dioxide that is annually sequestered as wood and/or vegetation by forestry"

Fourth, information on the actual level of provision of the PG in the area of interest is given:
"In the hilly and mountain areas of Emilia-Romagna, every year, 1.5 million tons of carbon are sequestered, corresponding to around 2.3 tons per hectare of forestry"

Fifth, the open-ended question to elicit the WTP is asked, as follows:
"How much is the maximum annual amount that your family is willingness to pay in order to reach the optimal level of sequestered carbon capacity in the hilly and mountain areas?"

Sixth, if the respondent is willing to pay, he/she is asked to express with respect to the previous payment his preference in terms of payment modality, choosing among:

- Generic private body
- Public authority ad hoc
- Increase of current taxation

Table A.1
Summary statistics of the samples' main characteristics in Application 1.

Socio-demographic characteristics	Study site: provinces of Emilia-Romagna (excluded the province of Bologna)	Policy site: Bologna province
Samples dimension	717	290
Mean WTP for soil erosion[a]	60.58 (0-40-300)	64.97 (0-45-300)
Mean WTP for carbon sequestration[a]	54.57 (0-30-300)	61.33 (0-30-300)
Average age	41.88 (18-41-99)	41.49 (18-40-77)
Share of male (%)	52.99	45.17
Average household size	2.94	2.84
Average number of minor	0.74	0.59
Average number of elderly	0.31	0.30
Share of households with farmer (%)	11.44	14.83
Share of unemployed (%)	36.54	32.07
Share of university degree (%)	36.27	42.07
Level of education (%)		
1 – primary school	0.69	0.34
2 – secondary school	10.32	8.28
3 – higher school	52.72	50.34
4 – BA degree	14.23	13.45
5 – MA degree	20.92	26.55
6 – other postgraduate	1.12	1.03
Average monthly household income[a]	2771 (1000–2000-8000)	2900 (1000–3000-8000)
Average annual payment for LRA[b]	124.63 (0-30-10,000)	140.50 (0-30-10,000)
Average annual payment for food basket[a]	2679 (10–2000-10,000)	2772 (10–2000-10,000)
Residence in hill-mountain areas (%)	19.94	17.93

[a]: values expressed in Euro.
[b]: LRA -Land Reclamation Authority-. Values in parentheses are minimum, median and maximum.

Table A.2
Summary statistics of the samples' main characteristics in Application 2.

Socio-demographic characteristics	Study site: provinces of Emilia-Romagna (excluded the province of Ferrara)	Policy site: Ferrara province
Samples dimension	939	68
Mean WTP for soil erosion[a]	62.68 (0-40-300)	50.26 (0-30-300)
Mean WTP for carbon sequestration[a]	57.24 (0-30-300)	46.54 (0-22-300)
Average age	41.79 (18-41-99)	41.41 (19-41-71)
Share of male (%)	51.01	47.06
Average household size	2.91	2.93
Average number of minor	0.69	0.60
Average number of elderly	0.30	0.32
Share of households with farmer (%)	12.14	16.18
Share of unemployed (%)	34.93	39.71
Share of university degree (%)	38.66	29.41
Level of education (%)		
1 – primary school	0.53	1.47
2 – secondary school	9.15	17.65
3 – higher school	52.08	51.47
4 – BA degree	13.74	17.64
5 – MA degree	23.32	11.76
6 – other postgraduate	1.17	0
Average monthly household income[a]	2831 (1000–3000-8000)	2500 (1000–2000-8000)
Average annual payment for LRA[b]	132.30 (0-30-10,000)	85.13 (0-42.50-1200)
Average annual payment for food basket[a]	2727 (10–2000-10,000)	2412 (10–1800-10,000)
Residence in hill-mountain areas (%)	20.45	4.41

[a]: values expressed in Euro.
[b]: LRA -Land Reclamation Authority-. Values in parentheses are minimum, median and maximum.

Table A.3
Variables description.

Variable	Coding	Modalities
Residence in hilly and mountain areas	*res*	0 = no; 1 = yes
Residence in the province capital	*admin_centre*	0 = no; 1 = yes
Ownership of a house in hilly and mountain areas	*mount_house*	0 = no; 1 = yes
Ownership of agricultural land in hilly and mountain areas	*mount_land*	0 = no; 1 = yes
Having relatives in hilly and mountain areas	*mount_rela*	0 = no; 1 = yes
Age	*age*	
Age squared	*age_squared*	
Gender	*gender*	0 = male; 1 = female
Being employed	*employ*	0 = no; 1 = yes
Educational level	*edu_level*	1 = primary school; 2 = secondary school; 3 = higher school; 4 = BA degree; 5 = MA degree; 6 = other postgraduate
Having a university degree	*univ_degree*	0 = no; 1 = yes

(*continued on next page*)

Table A.3 (*continued*)

Variable	Coding	Modalities
Number of household members	*household_size*	
Numbers of household minor	*household_minor*	
Numbers of household elderly	*household_eld*	
Presence of farmer(s) in the household	*household_farmer*	0 = no; 1 = yes
Household average annual payment for a food basket[a]	*basket_pay*	
Household average annual payment for Land Reclamation Authority (LRA)[a]	*lra_pay*	
Household average monthly income[a]	*household_income*	
Soil erosion is considered relevant	*rel_eros*	0 = no; 1 = yes
Carbon sequestration is considered relevant	*rel_ carb*	0 = no; 1 = yes
Rural vitality is considered relevant	*rel_ rurvit*	0 = no; 1 = yes
Respondent is indifferent to soil erosion	*rel_eros_idk*	0 = no; 1 = yes
Respondent is indifferent to carbon sequestration	*rel_carb_idk*	0 = no; 1 = yes
Respondent is indifferent to rural vitality	*rel_rurvit_idk*	0 = no; 1 = yes
Land abandonment impacts positively on soil erosion	*aband_eros_p*	0 = no; 1 = yes
Land abandonment impacts positively on carbon sequestration	*aband_carb_p*	0 = no; 1 = yes
Land abandonment impacts negatively on soil erosion	*aband_eros_n*	0 = no; 1 = yes
Land abandonment impacts negatively on carbon sequestration	*aband_carb_n*	0 = no; 1 = yes
Land abandonment impacts positively on rural vitality	*aband_rurvit_p*	0 = no; 1 = yes
Land abandonment impacts negatively on rural vitality	*aband_rurvit_n*	0 = no; 1 = yes
Modality of WTP payment	*how_wtp_pay*	0 = no payment; 1 = private body; 2 = public authority ad hoc; 3 = increasing of current taxation
Practicing fishing	*fishing*	0 = no; 1 = yes
Practicing hunting	*hunting*	0 = no; 1 = yes
Going to food festivals and other food-related recreational activities	*recreation*	0 = no; 1 = yes

[a]*: values expressed in Euro.*

Appendix B. Function transfer models

Equation (B.1) depicts the multiple linear regression model for soil erosion:

$$\begin{aligned} wtp\_eros = {} & \beta_0 + \beta_1 \cdot mount\_house + \beta_2 \cdot univ\_degree + \beta_3 \cdot household\_farmer + \beta_4 \cdot lra\_pay + \\ & + \beta_5 \cdot household\_income + \beta_6 \cdot rel\_rurvit + \beta_7 \cdot rel\_(eros\_idk) + \beta_8 \cdot aband\_eros\_p + \\ & + \beta_9 \cdot aband\_rurvit\_p + \beta_{10} \cdot aband\_eros\_n + \beta_{11} \cdot how\_wtp\_pay + \beta_{12} \cdot fishing + \\ & + \beta_{13} \cdot hunting + \beta_{14} \cdot recreation \end{aligned} \tag{B.1}$$

Table B.1
Results of the multiple linear regression model on the WTP for soil erosion (wtp_eros) in applications 1 and 2.

	Application 1 (policy site: Bologna province; study site: other provinces)					Application 2 (policy site: Ferrara province; study site: other provinces)				
Variables	Coeff.	S. E.	t value	Pr (>|t|)	Sign.	Coeff.	S. E.	t value	Pr (>|t|)	Sign.
Intercept	5.564	15.030	0.370	0.711		0.989	13.084	0.076	0.939	
mount_house: 1	9.797	5.346	1.832	0.067	.	7.546	4.694	1.608	0.108	
univ_degree: 1	−12.440	5.117	−2.431	0.015	*	−9.080	4.465	−2.033	0.042	*
household_farmer: 0	−31.072	7.855	−3.956	8.41e-05	***	−22.023	6.878	−3.202	0.001	**
lra_pay	0.026	0.004	6.750	3.11e-11	***	0.028	0.003	8.217	7.05e-16	***
household_income	0.005	0.002	3.124	0.002	**	0.006	0.002	3.762	0.001	***
rel_rurvit: 1	20.669	8.759	2.360	0.019	*	26.111	7.591	3.440	0.001	***
rel_eros_idk: 1	−14.358	6.234	−2.303	0.022	*	−13.797	5.623	−2.454	0.014	*
aband_eros_p: 1	21.734	6.923	3.140	0.002	**	20.965	6.182	3.391	0.001	***
aband_rurvit_p: 1	−15.792	7.945	−1.988	0.047	*	−21.953	6.873	−3.194	0.001	**
aband_eros_n: 1	13.953	6.394	2.182	0.029	*	10.794	5.675	1.902	0.057	.
how_wtp_pay: 1	65.503	10.681	6.132	1.45e-09	***	61.617	9.539	6.459	1.70e-10	***
how_wtp_pay: 2	56.069	8.046	6.968	7.42e-12	***	58.513	7.242	8.079	2.03e-15	***
how_wtp_pay: 3	43.726	8.679	5.038	5.98e-07	***	47.347	7.824	6.052	2.08e-09	***
fishing: 1	29.907	9.703	3.082	0.002	**	30.867	8.675	3.558	0.001	***
hunting: 1	20.319	8.505	2.389	0.017	*	20.098	7.832	2.566	0.010	*
recreation: 1	11.144	5.056	2.204	0.028	*	5.640	4.516	1.249	0.212	
pseudo-R^2	0.317					0.315				

Variables coding: mount_house, 0 = no, 1 = yes; univ_degree, 0 = no, 1 = yes; household_farmer, 0 = no, 1 = yes; rel_rurvit, 0 = no, 1 = yes; aband_" …"_" …", 0 = no, 1 = yes; how_wtp_pay, 0 = no, 1 = private body, 2 = public authority ad hoc, 3 = increased taxation; fishing, 0 = no, 1 = yes; hunting, 0 = no, 1 = yes; recreation, 0 = no, 1 = yes. Coeff.: coefficients. S.E.: standard error. t value: test statistic. Pr(>|t|): p-value. Sign.: significance levels, '***' = 0.001; '**' = 0.01; '*' = 0.05; '.' = 0.1.

Equation (B.2) depicts the multiple linear regression model for carbon sequestration:

$$\begin{aligned} wtp\_carb = {} & \beta_0 + \beta_1 \cdot mount\_house + \beta_2 \cdot age\_squared + \beta_3 \cdot employ + \beta_4 \cdot univ\_degree + \\ & + \beta_5 \cdot household\_eld + \beta_6 \cdot household\_farmer + \beta_7 \cdot lra\_pay + \beta_8 \cdot household\_income + \\ & + \beta_9 \cdot rel\_carb + \beta_{10} \cdot rel\_eros\_idk + \beta_{11} \cdot aband\_eros\_p + \beta_{12} \cdot aband\_eros\_n + \\ & + \beta_{13} \cdot how\_wtp\_pay + \beta_{14} \cdot fishing + \beta_{15} \cdot hunting \end{aligned} \tag{B.2}$$

Table B.2
Results of the multiple linear regression model on the WTP for carbon sequestration (wtp_carb) in applications 1 and 2.

	Application 1 (policy site: Bologna province; study site: other provinces)					Application 2 (policy site: Ferrara province; study site: other provinces)				
Variables	Coeff.	S. E.	t value	Pr (>|t|)	Sign.	Coeff.	S. E.	t value	Pr (>|t|)	Sign.
Intercept	4.058	13.251	0.306	0.759		5.411	11.970	0.452	0.651	
mount_house: 1	8.651	4.784	1.809	0.071	.	4.185	4.344	0.963	0.336	
age_squared	−0.004	0.002	−2.004	0.045	*	−0.002	0.002	−1.022	0.307	
employ: 1	9.008	4.724	1.907	0.057	.	12.727	4.395	2.896	0.004	**
univ_degree: 1	−11.968	4.788	−2.499	0.013	*	−5.997	4.303	−1.394	0.164	
household_elderly	9.305	3.685	2.525	0.012	*	6.172	3.512	1.758	0.079	.
household_farmer: 0	−24.025	7.391	−3.251	0.001	**	−19.351	6.607	−2.929	0.003	**
lra_pay	0.032	0.003	9.298	<2e-16	***	0.030	0.003	10.033	<2e-16	***
household_income	0.004	0.002	2.383	0.017	**	0.004	0.001	2.722	0.007	**
rel_carb: 1	10.485	5.563	1.885	0.019	.	11.241	5.080	2.213	0.027	*
rel_eros_idk: 1	−14.168	5.714	−2.480	0.013	*	−13.5690	5.284	−2.568	0.010	*
aband_eros_p: 1	23.421	6.401	3.659	0.000	**	20.871	5.879	3.550	0.000	***
aband_eros_n: 1	13.872	5.938	2.336	0.019	*	8.798	5.419	1.624	0.105	
how_wtp_pay: 1	46.783	9.908	4.722	2.83e-06	***	45.587	9.091	5.014	6.39e-07	***
how_wtp_pay: 2	47.091	7.410	6.355	3.76e-10	***	49.773	6.868	7.247	9.02e-13	***
how_wtp_pay: 3	40.456	7.957	5.084	4.74e-07	***	41.923	7.389	5.673	1.88e-08	***
fishing: 1	36.462	8.865	4.113	4.37e-05	***	44.349	8.131	5.454	6.32e-08	***
hunting: 1	18.878	7.810	2.417	0.016	*	15.449	7.395	2.089	0.037	*
pseudo-R^2	0.351					0.346				

Variables coding: mount_house, 0 = no, 1 = yes; univ_degree, 0 = no, 1 = yes; household_farmer, 0 = no, 1 = yes; rel_rurvit, 0 = no, 1 = yes; aband_" …"_" …", 0 = no, 1 = yes; how_wtp_pay, 0 = no, 1 = private body, 2 = public authority ad hoc, 3 = increased taxation; fishing, 0 = no, 1 = yes; hunting, 0 = no, 1 = yes; recreation, 0 = no, 1 = yes. Coeff.: coefficients. S.E.: standard error. T value: test statistic. Pr(>|t|): p-value. Sign.: significance levels, '***' = 0.001; '**' = 0.01; '*' = 0.05; '.' = 0.1.

Equation (B.3) depicts the Tobit model for soil erosion:

$$\begin{aligned} wtpc\_eros = {} & \beta_0 + \beta_1 \cdot age + \beta_2 \cdot age\_squared + \beta_3 \cdot employ + \beta_4 \cdot univ\_degree + \beta_5 \cdot household\_size + \beta_6 \cdot household\_farmer + \beta_7 \cdot lra\_pay + \beta_8 \cdot household\_income \\ & + \beta_9 \cdot rel\_eros\_idk \end{aligned} \tag{B.3}$$

Table B.3
Results of the censored (Tobit) regression model for soil erosion (wtp_eros) in applications 1 and 2.

	Application 1 (policy site: Bologna province; study site: other provinces)					Application 2 (policy site: Ferrara province; study site: other provinces)				
Variables	Coeff.	S. E.	z value	Pr (>|z|)	Sign.	Coeff.	S. E.	z value	Pr (>|z|)	Sign.
Intercept	139.358	27.980	4.981	6.34e-07	***	125.837	24.194	5.201	1.98e-07	***
age	−2.393	1.205	−1.987	0.047	*	0.027	0.012	2.335	0.019	*
age_squared	0.023	0.013	1.737	0.082	.	−2.676	1.055	0.011	0.082	*
employ: 1	4.948	6.580	0.752	0.452		16.324	5.913	2.761	0.006	**
univ_degree: 1	−10.750	6.304	−1.705	0.088	.	−9.032	5.416	−1.668	0.095	.
household_size	4.335	2.708	1.601	0.109		4.121	2.345	1.757	0.079	.
household_farmer: 0	−46.771	9.426	−4.962	6.97e-07	***	−37.149	8.068	−4.604	4.14e-06	***
lra_pay	0.028	0.005	6.213	5.20e-10	***	0.028	0.004	7.445	9.71e-14	***
household_income	0.007	0.002	3.137	0.002	**	0.008	0.002	4.134	3.56e-05	***
rel_erosion_idk: 1	−27.650	7.465	−3.704	0.000	***	−28.090	6.498	−4.323	1.54e-05	***
Log (scale)	4.319	0.029	149.074	<2e-16	***	4.324	0.025	171.992	<2e-16	***
pseudo-R^2	0.167					0.183				

Variables coding: employ, 0 = no, 1 = yes; univ_degree, 0 = no, 1 = yes; household_farmer, 0 = no, 1 = yes; rel_eros_idk, 0 = no, 1 = yes. Coeff.: coefficients. S.E.: standard error. z value: test statistic. Pr(>|z|): p-value. Sign.: significance levels, '***' = 0.001; '**' = 0.01; '*' = 0.05; '.' = 0.1.

Equation (B.4) depicts the Tobit model for carbon sequestration:

$$wtpc\_carb = \beta_0 + \beta_1 \cdot household\_farmer + \beta_2 \cdot lra\_pay + \beta_3 \cdot household\_income + \beta_4 \cdot rel\_eros\_idk \tag{B.4}$$

Table B.4
Results of the censored (Tobit) regression model for carbon sequestration (wtp_carb) in applications 1 and 2.

	Application 1 (policy site: Bologna province; study site: other provinces)					Application 2 (policy site: Ferrara province; study site: other provinces)				
Variables	Coeff.	S. E.	z value	Pr (>|z|)	Sign.	Coeff.	S. E.	z value	Pr (>|z|)	Sign.
Intercept	93.552	11.452	8.169	3.11e-16	***	84.774	10.204	8.308	<2.e−16	***
household_farmer: 0	−47.529	8.248	−5.763	8.28e-09	***	−41.724	7.460	−5.593	2.23e-08	***
Lra_pay	0.036	0.004	8.766	<2e-16	***	0.033	0.004	9.078	<2e-16	***
household_income	0.005	0.002	2.976	0.003	**	0.008	0.002	4.517	6.26e-06	***
Rel_erosion_idk: 1	−27.024	6.784	−3.983	6.80e-05	***	−27.376	6.178	−4.431	9.37e-06	***
Log (scale)	4.233	0.029	146.576	<2e-16	***	4.276	0.025	170.254	<2e-16	***
pseudo-R^2	0.201					0.205				

Variables coding: employ, 0 = no, 1 = yes; univ_degree, 0 = no, 1 = yes; household_farmer, 0 = no, 1 = yes; rel_eros_idk, 0 = no, 1 = yes. Coeff.: coefficients. S.E.: standard error. z value: test statistic. Pr(>|z|): p-value. Sign.: significance levels, '***' = 0.001; '**' = 0.01; '*' = 0.05; '.' = 0.1.

Appendix C. Empirical bounds of transferred distributions

The following figures depict the results of the leave-*k*-out strategy (with *k* equals to ¼ of the original study site/donor sample, in 1000 random draws) applied to determine the empirical bounds of the transferred distributions with respect to both the PGs under analysis in both Application 1 and Application 2.

All the results depicted by Figs. C1-C3 are confirmed by simulation studies carried out by dropping out the 10% of the original study site/donor sample in 1000 random draws, as well as by dropping out both the 10% and 25% in 5000 random draws. These results are contained in an appendix available upon request.

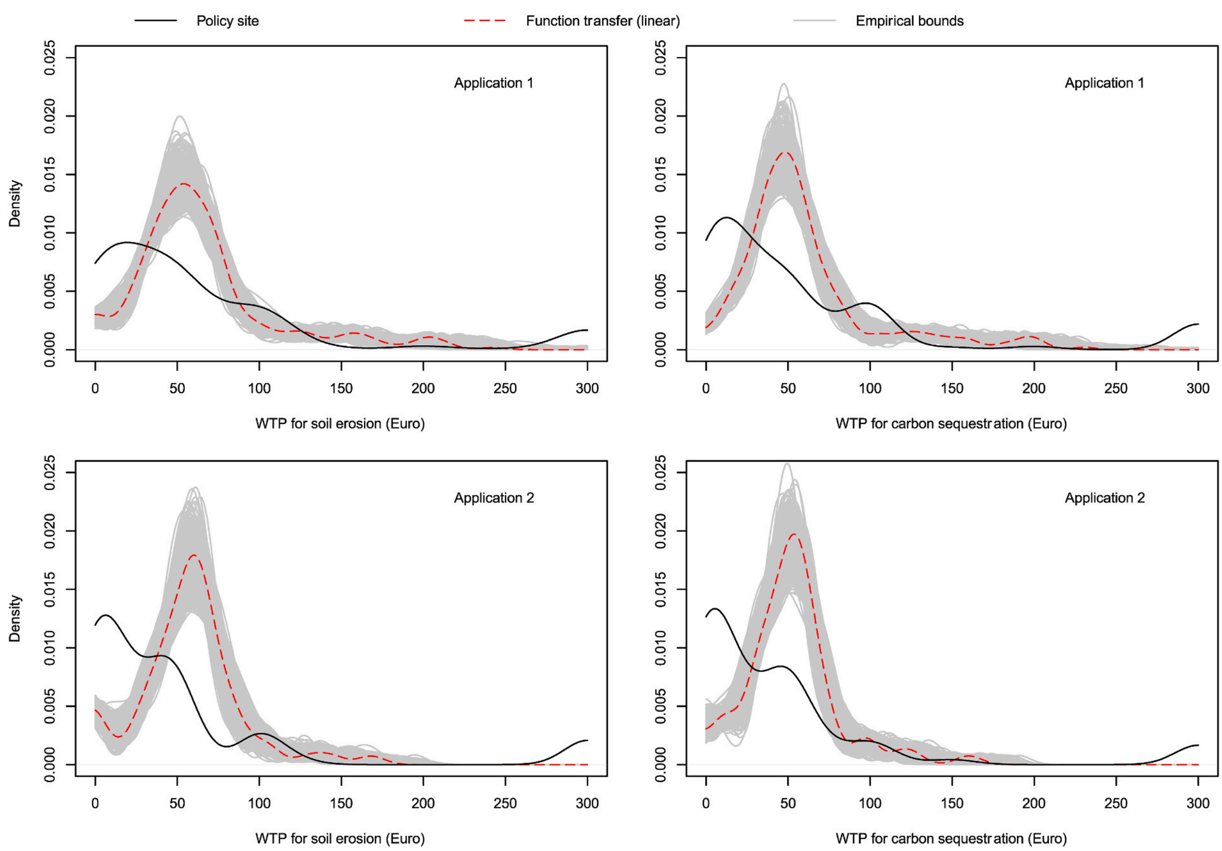

Fig. C1. Empirical bounds of transferred WTP distributions for soil erosion and carbon sequestration obtained in applications 1 and 2 by means of the linear function transfer.

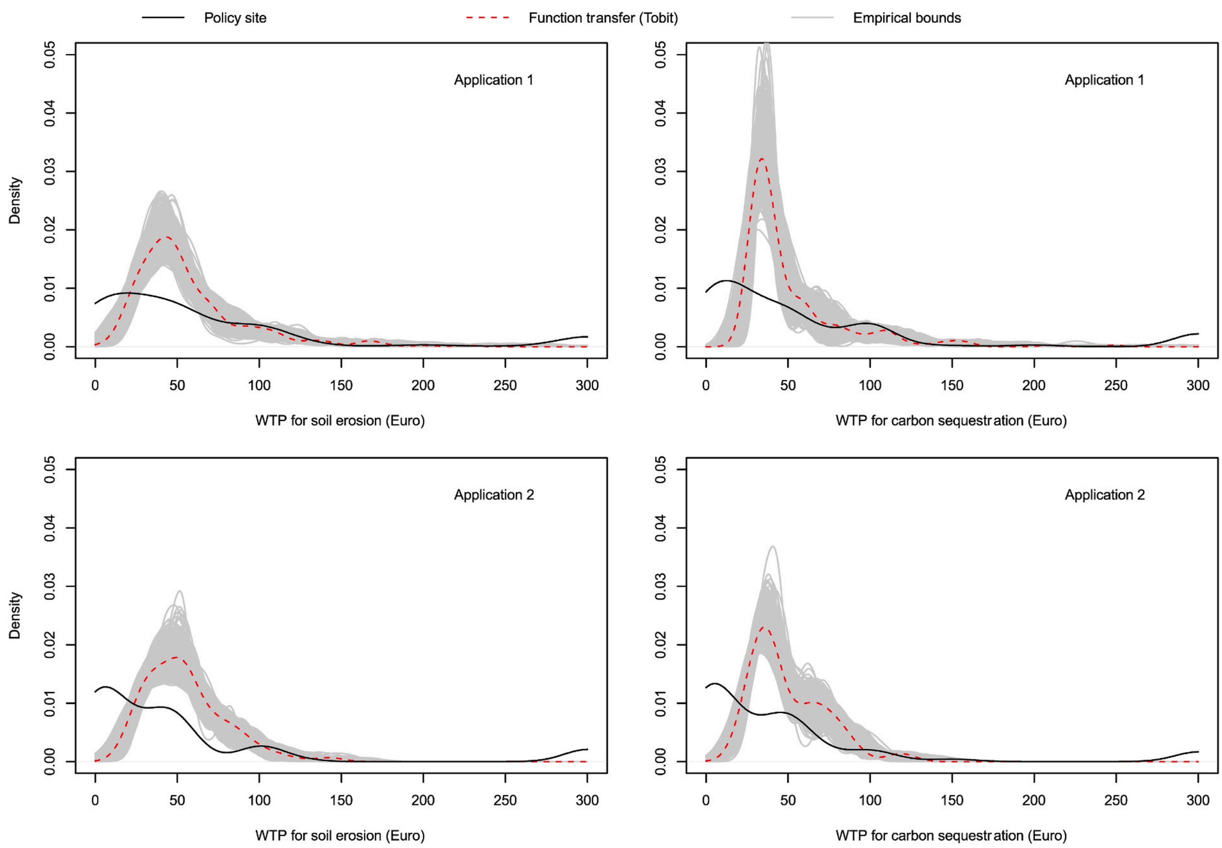

Fig. C2. Empirical bounds of transferred WTP distributions for soil erosion and carbon sequestration obtained in applications 1 and 2 by means of the censored (Tobit) function transfer.

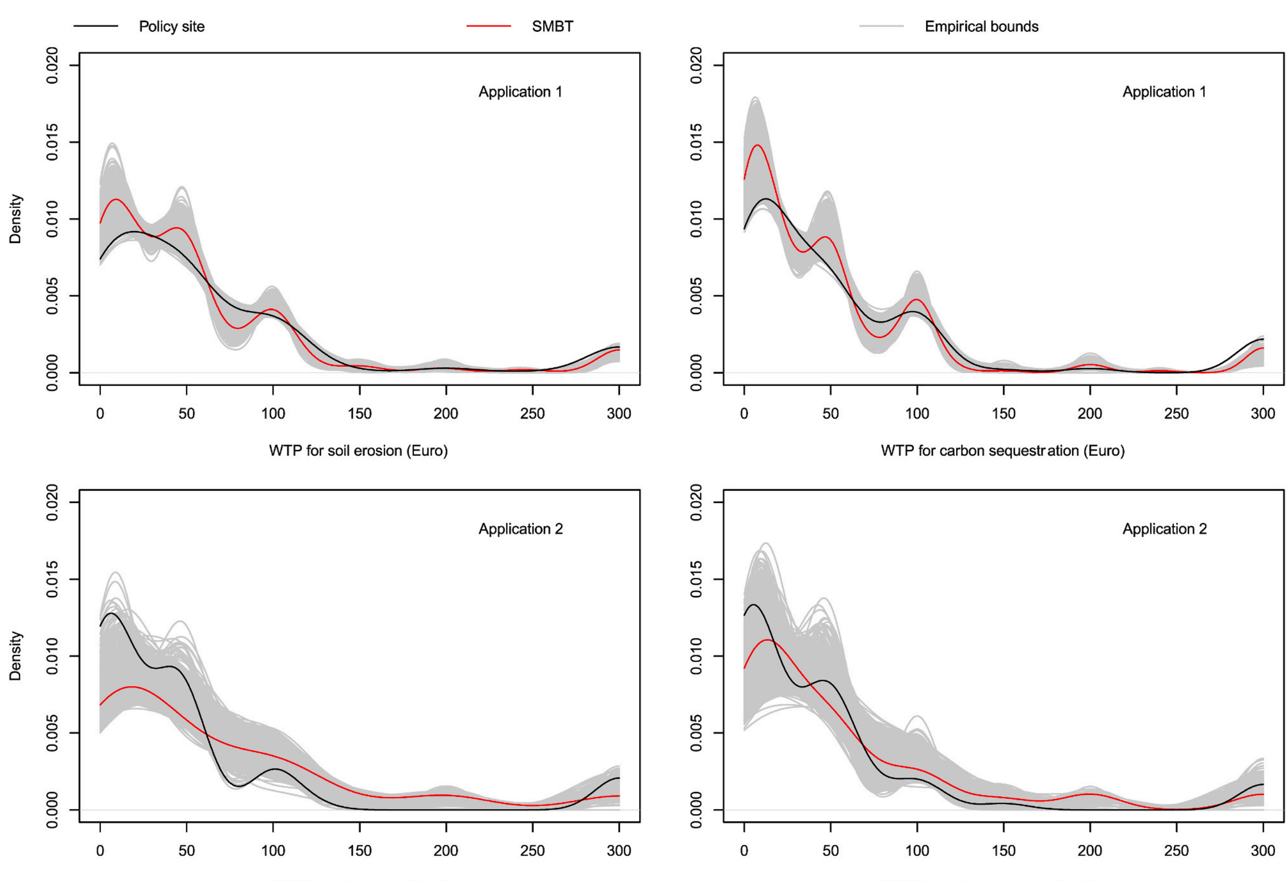

Fig. C3. Empirical bounds of transferred WTP distributions for soil erosion and carbon sequestration obtained in applications 1 and 2 by means of the Statistical Matching Benefit Transfer (SMBT).

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.seps.2020.100935.

Funding

This work was supported by European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program through the project PROVIDE, Grant Agreement No. 633838. This work does not necessarily reflect the view of the EU and in no way anticipates the Commission's future policy.

References

[1] Kleijn D, Sutherland WJ. How effective are European agri-environment schemes in conserving and promoting biodiversity? J Appl Ecol 2003;40:947–69. https://doi.org/10.1111/j.1365-2664.2003.00868.x.

[2] Lefebvre M, Espinosa M, Gomez y Paloma S, Paracchini ML, Piorr A, Zasada I. Agricultural landscapes as multi-scale public good and the role of the Common Agricultural Policy. J Environ Plann Manag 2015;58:2088–112. https://doi.org/10.1080/09640568.2014.891975.

[3] Pavlis ES, Terkenli TS, Kristensen SBP, Busck AG, Cosor GL. Patterns of agri-environmental scheme participation in Europe: indicative trends from selected case studies. Land Use Pol 2016;57:800–12. https://doi.org/10.1016/j.landusepol.2015.09.024.

[4] European Commission. COM/2018/392 final – 2018/0216 (COD). *Proposal for a regulation OF the EUROPEAN PARLIAMENT and OF the COUNCIL establishing rules on support for strategic plans to be drawn up by member States under the Common agricultural policy (CAP strategic plans) and financed by the European agricultural guarantee fund (eagf) and by the European agricultural Fund for rural development (EAFRD) and repealing reg*ulation (EU) No 1305/2013 of the European parliament and of the Council and regulation (EU) No 1307/2013 of the European parliament and of the Council. 2018.

[5] Bateman IJ, Carson R, Brett D, Hanemann M, Hanley N, Hett T, Lee MJ, Loomes G, Mourato S, Özdemiroglu E, Pearce DW. Economic valuation with stated preference techniques: a manual. Cheltenham: Edward Elgar Publishing; 2002.

[6] Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC. The benefit-transfer challenges. Annu Rev Resour Econ 2010;2:161–82. https://doi.org/10.1146/annurev.resource.012809.103933.

[7] Venkatachalam L. The contingent valuation method: a review. Environ Impact Assess Rev 2004;24:89–124. https://doi.org/10.1016/S0195-9255(03)00138-0.

[8] Louviere JJ, Hensher DA, Swait JD. Stated choice methods: analysis and applications. Cambridge: University Press; 2000.

[9] Hoyos D. The state of the art of environmental valuation with discrete choice experiments. Ecol Econ 2010;69:1595–603. https://doi.org/10.1016/j.ecolecon.2010.04.011.

[10] Hanley N, Wright RE, Alvarez-Farizo B. Estimating the economic value of improvements in river ecology using choice experiments: an application to the water framework directive. J Environ Manag 2006;78:183–93. https://doi.org/10.1016/j.jenvman.2005.05.001.

[11] Collins R, Schaafsma M, Hudson MD. The value of green walls to urban biodiversity. Land Use Pol 2017;64:114–23. https://doi.org/10.1016/j.landusepol.2017.02.025.

[12] Dachary-Bernard J, Rambonilaza T. Choice experiment, multiple programmes contingent valuation and landscape preferences: how can we support the land use decision making process? Land Use Pol 2012;29:846–54. https://doi.org/10.1016/j.landusepol.2012.01.002.

[13] Johnston RJ, Rolfe J, Rosenberger R, Brouwer R, editors. Benefit transfer of environmental and resource values. Heidelberg: Springer Netherlands; 2015.

[14] Richardson L, Loomis J, Kroeger T, Casey F. The role of benefit transfer in ecosystem service valuation. Ecol Econ 2015;115:51–8. https://doi.org/10.1016/j.ecolecon.2014.02.018.

[15] Bal F, Button KJ, Nijkamp P. Ceteris paribus, meta-analysis and value transfer. Soc Econ Plann Sci 2002;36:127–38. https://doi.org/10.1016/S0038-0121(01)00022-2.

[16] Fitzpatrick L, Parmeter CF, Agar J. Threshold effects in meta-analyses with application to benefit transfer for coral reef valuation. Ecol Econ 2017;133:74–85. https://doi.org/10.1016/j.ecolecon.2016.11.015.

[17] Czajkowski M, Ahtiainen H, Artell J, Meyerhoff J. Choosing a functional form for an international benefit transfer: evidence from a nine-country valuation experiment. Ecol Econ 2017;134:104–13. https://doi.org/10.1016/j.ecolecon.2017.01.005.

[18] Artell J, Ahtiainen H, Pouta E. Distance decay and regional statistics in international benefit transfer. Ecol Econ 2019;164:106–383. https://doi.org/10.1016/j.ecolecon.2019.106383.

[19] Lloyd-Smith P, Zawojska E, Adamowicz WL. Moving beyond the contingent valuation versus choice experiment debate – presentation effects in stated preference. Working Papers 2018-14. University of Warsaw; 2018.

[20] Morrison M, Bennett J, Blamey R, Louviere J. Choice modeling and tests of benefit transfer. Am J Agric Econ 2002;84:161–70. https://doi.org/10.1111/1467-8276.00250.

[21] Whitehead JC, Morgan OA, Huth WL. Benefit transfers with the contingent valuation method. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R, editors. Benefit transfer of environmental and resource values. Dordrecht: Springer Netherlands; 2015. p. 119–40. https://doi.org/10.1007/978-94-017-9930-0_7.

[22] Johnston RJ, Boyle KJ, Adamowicz W, Bennett J, Brouwer R, Cameron TA, Hanemann WM, Hanley N, Ryan M, Scarpa R, Tourangeau R, Vossler CA. Contemporary guidance for stated preference studies. Journal of the Association of Environmental and Resource Economists 2017;4:319–405.

[23] Kaul S, Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC. What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. J Environ Econ Manag 2013;66:90–104. https://doi.org/10.1016/j.jeem.2013.03.001.

[24] Newbold S, Simpson RD, Massey DM, Heberling MT, Wheeler W, Corona J, Hewitt J. Benefit transfer challenges: perspectives from U.S. Practitioners. Environ Resour Econ 2018;69:467–81. https://doi.org/10.1007/s10640-017-0207-7.

[25] Rosenberger RS. Benefit transfer validity and reliability. In: Johnston RJ, Rolfe J, Rosenberger RS, Brouwer R, editors. Benefit transfer of environmental and resource values. Dordrecht: Springer Netherlands; 2015. p. 307–26. https://doi.org/10.1007/978-94-017-9930-0_14.

[26] Conti PL, Marella D, Scanu M. How far from identifiability? A systematic overview of the statistical matching problem in a non parametric framework. Commun Stat Theor Methods 2017;46:967–94. https://doi.org/10.1080/03610926.2015.1010005.

[27] Singh AC, Mantel H, Kinack M, Rowe G. Statistical matching: use of auxiliary information as an alternative to the conditional independence assumption. Surv Methodol 1993;19:59–79.

[28] Denk M, Hackl P. Data integration and record matching: an Austrian contribution to research in official statistics. Aust J Stat 2003;32:305–21. https://doi.org/10.17713/ajs.v32i4.464.

[29] Donatiello G, D'Orazio M, Frattarola D, Rizzi A, Scanu M. Statistical matching of income and consumption expenditures. Int J Econ Sci 2014;16:50–65.

[30] Ballin M, D'Orazio M, Zio MD, Scanu M, Torelli N. Statistical matching of two surveys with a common subset. Working paper n. 124. Università degli Studi di Trieste; 2009.

[31] Roesch A, Lips M. Sampling design for two combined samples of the farm accountancy data Network (FADN). J Agric Biol Environ Stat 2013;18:178–203. https://doi.org/10.1007/s13253-013-0130-5.

[32] D'Alberto R, Zavalloni M, Raggi M, Viaggi D. AES impact evaluation with integrated farm data: combining statistical matching and propensity score matching. Sustainability 2018;10:1–24. https://doi.org/10.3390/su10114320.

[33] Poppe KJ, Wolfert J, Verdouw CN, Renwick A. A European perspective on the economics of big data. Farm Policy Journal 2015;12:11–9. ISSN 1449-8812.

[34] UNECE, United Nations Economic Commission for Europe – Statistics Division. High level group for the modernization of official statistics (HLG MOS strategy). 2020. https://www.unece.org/stats/stats_h.html [accessed 06 April 2020].

[35] EC H2020. European Commission Horizon2020 Research and Innovation Programme; 2020. https://ec.europa.eu/programmes/horizon2020/en [accessed 14 May 2020].

[36] Norton D, Hynes S. Estimating the benefits of the Marine Strategy Framework Directive in Atlantic member states: a spatial value transfer approach. Ecol Econ 2018;151:82–94. https://doi.org/10.1016/j.ecolecon.2018.04.024.

[37] Bateman IJ, Brouwer R, Ferrini S, Schaafsma M, Barton DN, Dubgaard A, Hasler B, Hime S, Liekens I, Navrud S, De Nocker L, Ščeponavičiūtė R, Semėnienė D. 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. Environ Resour Econ 2011;50:365–87. https://doi.org/10.1007/s10640-011-9476-8.

[38] Raggi M, Ronchi D, Viaggi D. Valutazione economica del miglioramento qualitativo della risorsa idrica: un'applicazione di benefit transfer al bacino del Po. In: Casini L, Gallerani V, Viaggi D, editors. Acqua, agricoltura e ambiente nei nuovi scenari di Politica Comunitaria. Milano: Franco Angeli; 2008. p. 49–65. ISBN 978-88-568-0.

[39] Newbold S, Walsh PJ, Massey DM, Hewitt J. Using structural restrictions to achieve theoretical consistency in benefit transfers. Environ Resour Econ 2018;69:529–53. https://doi.org/10.1007/s10640-017-0209-5.

[40] Johnston RJ, Rosenberger RS. Methods, trends and controversies in contemporary benefit transfer. J Econ Surv 2009;24:479–510. https://doi.org/10.1111/j.1467-6419.2009.00592.x.

[41] Brouwer R, Martin-Ortega J, Dekker T, Sardonini L, Andreu J, Kontogianni A, Skourtos M, Raggi M, Viaggi D, Pulido-Velazquez M, Rolfe J, Windle J. Improving value transfer with socio-economic adjustments in a multicountry choice experiment of water conservation alternatives. Aust J Agric Resour Econ 2015;59:1–21. https://doi.org/10.1111/1467-8489.12099.

[42] D'Orazio M, Di Zio MD, Scanu M. Statistical matching: theory and practice. Chichester: John Wiley & Sons; 2006.

[43] Mardia K, Kent J, Bibby J. Multivariate analysis. Cambridge: Academic Press; 1979.

[44] D'Alberto R, Raggi M. How much reliable are the integrated 'live' data? A validation strategy proposal for the non-parametric micro statistical matching. J Appl Stat 2020:1–27. https://doi.org/10.1080/02664763.2020.1724272. in press.

[45] Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. J Econ Surv 2008;22:31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x.

[46] D'Orazio M. Statistical matching and imputation of survey data with StatMatch. CRAN – R Software package; 2015.

[47] Harrell F. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag; 2001.

[48] Khan I, Zhao M, Ullah Khan S, Yao L, Ullah A, Xu T. Spatial heterogeneity of preferences for improvements in river basin ecosystem services and its validity for benefit transfer. Ecol Indicat 2018;98:627–37. https://doi.org/10.1016/j.ecolind.2018.05.018.

[49] Fitzpatrick L, Parmeter CF, Agar J. Approaches for visualizing uncertainty in benefit transfer from metaregression. Ecol Econ 2019;164:166–343. https://doi.org/10.1016/j.ecolecon.2019.05.024.

[50] Markatou M, Chen Y, Afendras G, Lindsay BG. Statistical distances and their role in robustness. In: Chen D-G, Jin Z, Li G, Li Y, Liu A, Zhao Y, editors. New advances in statistics and data science. Cham: Springer International Publishing; 2017. p. 3–26. https://doi.org/10.1007/978-3-319-69416-0_1.

[51] Gulati A, Rai SC. Farmers' willingness-to-pay towards soil and water conservation measures in agro-ecosystems of Chotanagpur Plateau, India. Water Environ J 2015; 29:523–32. https://doi.org/10.1111/wej.12151.

[52] Colombo S, Hanley N, Calatrava-Requena J. Designing policy for reducing the off-farm effects of soil erosion using choice experiments. J Agric Econ 2005;56:81–95. https://doi.org/10.1111/j.1477-9552.2005.tb00123.x.

[53] Colombo S, Calatrava-Requena J, Hanley N. Analysing the social benefits of soil conservation measures using stated preference methods. Ecol Econ 2006;58: 850–61. https://doi.org/10.1016/j.ecolecon.2005.09.010.

[54] Ellingson L, Seidl A. Comparative analysis of non-market valuation techniques for the Eduardo Avaroa Reserve, Bolivia. Ecol Econ 2007;60:517–25. https://doi.org/10.1016/j.ecolecon.2006.07.014.

[55] Roberts RM, Jones KW, Seidl A, Ek A, Smith H. Conservation finance and sustainable tourism: the acceptability of conservation fees to support the Tambopata National Reserve, Peru. J Sustain Tourism 2017;25:1353–66. https://doi.org/10.1080/09669582.2016.1257630.

[56] Alavalapati JRR, Shrestha RK, Stainback GA, Matta JR. Agroforestry development: an environmental economic perspective. Agrofor Syst 2004;61:299–310. https://10.1023/B:AGFO.0000029006.64395.72.

[57] Manski CF. Communicating uncertainty in policy analysis. P Natl Acad Sci USA 2019;116:7634–41. https://doi.org/10.1073/pnas.1722389115.

[58] Raggi M, Viaggi D, Bartolini F, Furlan A. The role of policy priorities and targeting in the spatial location of participation in Agri-Environmental Schemes in Emilia-Romagna (Italy). Land Use Pol 2015;47:79–89. https://10.1016/j.landusepol.2015.03.005.

D'Alberto Riccardo is *Research Fellow;* he is a statistician and his main research interests are data integration, Statistical Matching imputation, causal-effects analysis, impact evaluation and counterfactual approach analysis.

Zavalloni Matteo is *Junior assistant professor (fixed-term);* he is an agricultural and environmental economist.

Raggi Meri is *Assistant professor*; she obtained the PhD in "Statistical methodology for scientific research" in 2002 at the Department of Statistical Sciences of the University of Bologna, with a dissertation entitled "Asymptotic properties of Cross-Validation methods in regression problems". She is the co-author of over 150 scientific articles, (57 publications –H-index = 13– listed in SCOPUS; 41 listed in ISI Thomson), more than 200 conference presentations, and a deep experience with EU financed projects (20 in total; 1 in FP5, 3 in FP6, 9 in FP7 and 7 in H2020).

Viaggi Davide is currently *Professor*; he participates in research projects funded by the European Union, the Ministry of University and Research, Region Emilia-Romagna. Among them, he coordinated 3 EU projects (CAP-IRE and CLAIM of FP7, PROVIDE in H2020) and participated in about other 20. He is currently coordinating a 4th H2020 Project, CONSOLE. The main research themes concern the common agricultural policy (CAP), the bioeconomy, the economics and management of water resources and the evaluation of agri-environmental policies. Main economic tools used include linear programming, principal-agent models, multicriteria analysis, cost-benefits analysis and externality accounting and evaluation. The results have been presented in about 300 publications. He is currently president of the Italian Association of Agricultural and Applied Economics.