A META-ANALYSIS OF ECOSYSTEM SSERVICES ASSOCIATED WITH WETLANDS IN USFWS NATIONAL WILDLIFE REFUGES

Douglas Patton, John Bergstrom, Rebecca Moore, and Alan Covich

AUTHORS: University of Georgia, Athens, GA 30602
REFERENCE: Proceedings of the 2013 Georgia Water Resources Conference, *held April 10–11, 2013, at the University of Georgia*

Abstract. We conduct a novel meta-analysis of ecosystem service economic valuation results restricted to wetlands of the contiguous US. We focus our analysis on the estimation of benefits that wetlands contribute to water quality and flood and storm protection services. We developed a new statistical approach to enhance the accuracy and flexibility of our meta-analysis while maintaining a systematic approach. Using the new modeling method and the new domestic-oriented wetland valuation database, we estimate ecosystem service values at four case study National Wildlife Refuges.

INTRODUCTION

In recent years, public benefits from ecosystem services have been the focus of increasing attention from scientists, managers, and stakeholders. The measurement of the non-market benefits such as clean water or flood protection that the public receives from ecosystem services is known as ecosystem service valuation (Champ, Boyle, and Brown 2003). Primary valuation studies use original data to measure these values and are the ideal approach. Benefit transfer (BT) methods have been developed as a low-cost and rapid alternative to primary valuation studies. Benefit transfers are methods where ecosystem service values that have been previously measured by primary valuation studies are used individually or in a group to infer the value of benefits at an unstudied site. Meta-analysis benefit transfer (MABT) is a robust type of benefit transfer that typically uses a statistical model to incorporate information from multiple primary valuation studies to make a single benefit transfer.

The focus of our research is to develop improved techniques for conducting meta-analysis benefit transfer in order to gain a better understanding of the ecosystem service values associated with wetlands in the USFWS National Wildlife Refuge (NWR) System. A wide variety of policy applications ranging from low stakes local decisions to high stakes international policies require estimates of the benefits provided by various ecosystems. Benefit transfers are a valuation approach particularly apt for low-stakes decision making contexts that tend to exclude quantitative non-market ecosystem service values. Low-cost, rapid estimates of the economic value of wetland ecosystem services can be produced via benefit transfer with existing wetland ecosystem service meta-analysis (MA) models, but the accuracy of these models compared to primary studies is poorly understood. Compounding uncertainty about the appropriate uses of benefit transfer value estimates obtained from meta-analysis models is the uncertainty that exists with regard to the correct methodology for calculating benefit transfer estimates.

In this paper, we examine the flows of selected ecosystem services from wetlands in Arrowwood National Wildlife Refuge, North Dakota; Blackwater NWR, Maryland; Okefenokee NWR, Georgia; and Sevilleta and Bosque del Apache NWRs, New Mexico. The choice of sites is intended to contrast major types of wetlands of the contiguous United States. Our analysis focuses on two specific wetland-related ecosystem services: water quality provisioning and flood control (including storm surge protection).

For our analysis, we develop a novel MA database and implement two estimation strategies. The first strategy is the familiar OLS model and the second strategy is the PLWLS estimator. The unique MA database that we develop focuses entirely on wetlands in the contiguous United States. While we are most interested in flood control and water quality provisioning, we include a variety of other services and necessary controls in order to obtain better inference on parameters common across all observations. In important contribution of our MA model is that we normalize the dependent variable, aggregate willingness to pay, by both acres valued and the population over which the welfare measure is aggregated. We chose such a normalization as it became evident in the course of constructing the database that aggregating a welfare measure is too sensitive to potentially arbitrary assumptions to rely on a reduced form model. Specifically, we found that surface area and population were the most important aspects of the modeling process that are best treated as analyst imposed for the purpose of answering specific questions. An important implication of how we discuss and treat the dependent variable in the context of MABT is that we are forecasting the results of a primary valuation study, not that we are predicting a welfare measure that can be sensibly communicated without also describing the methodology by which the value was measured.

The appropriate interpretation of the dependent variable in benefit transfer models is an important and nuanced issue that in our view has received too little attention in the ecosystem service MABT literature. In all studies in this field that we are aware of, the dependent variable is treated directly as an estimate of a measure of a willingness to pay (WTP) construct. We interpret the dependent variable somewhat differently, as the resulting welfare estimate of a primary valuation study. For forecasts of the dependent variable the value is fundamentally the predicted result of a primary valuation study. Due to the possibility of generating predictions from unlikely or even nonsensical combinations of services and methods, such an interpretation better highlights the context of the WTP estimate and implies the appropriate means for criteria validation. For example, via MABT one might forecast the value estimated by a travel cost study that values commercial fishing at a site where no commercial fishing operations exist; criterion validation via a primary valuation study of this predicted value is not feasible. The simulation of the results of such a study are feasible but the validity of the results fails a basic test of content validity. Similarly, the practice of coding methodological variables at their sample means and predicting a single number suggests an unusual hybrid of studies that would defy criterion validation (Bishop 2003); alternatively, simulating a number of realistic primary valuation studies and using a weighted average to combine the results is more consistent with the original data and more amenable to criterion validation with future primary valuation studies.

In the sections below we discuss the two modeling methodologies we use, focusing on developing an intuitive undertanding of the latter model, a novel regression approach. Results and an appropriate discussion and conclusions follow with recommendations for future research.

METHODS

The OLS Regression

In order to allow comparison between our MA dataset and existing datasets, we provide an OLS regression that is largely consistent with the existing wetland meta-analyses. Diverging somewhat from existing wetland MAs, we utilize the dependent variable normalized by both surface area in acres and the population over which individual welfare estimates were aggregated in the primary valuation study. Equation (1) is the basic form for our linear MABT forecast model where the dependent variable, $\log(WTP_i)$ is the natural logarithm of the estimated aggregate willingness to pay divided by the population and acreage counts associated with the aggregate value for observation i.

$$\begin{aligned}\log(WTP_i) = \beta_1 + study\backslash site_{vars\ i} * \beta_2 \\ + methodological_{vars_i} * \beta_3 \\ + context_{vars_i} * \beta_4 + e_i\end{aligned} \tag{1}$$

In equation (1), *study/site*$_{vars}$ is a set of study descriptor variables, including the study area and the count of the study population over which the valuation is aggregated; both of these values are used in the normalization of the dependent variable. We specifically label this group of variables to draw attention to the fact that the acres and population over which valuation is applied are characteristics of a primary valuation study and not natural characteristics of a landscape or geographic context. *Methodological*$_{vars}$ includes dummy variables for the service valued and indicators of whether the valuation approach is stated preference, revealed preference, or replacement cost. We group the remaining variables into the broad category of *context*$_{vars}$, which includes variables that describe both the user population and landscape in the vicinity of the primary valuation study site. The final term, e_i, is assumed to be a mean zero independently distributed random variable.

We include in the last category of context variables a ratio of local, surrounding wetlands to a local population count as an indicator of the availability of substitutes outside the study area relative to the local population. These two variables are measured in acres and a count of people, but are distinct from the acreage and population counts in the variable group, *study/site*$_{var}$. Other contextual variables include the average GDP per capita of the one or more states in close proximity to the study site and indicators of wetland type.

An important feature of our meta-analysis model is the normalization we chose for the dependent variable. Our choice was motivated by our observation that most primary valuation studies use effectively obscure reasoning for choosing the population size for aggregating welfare measures and the boundaries and types of land included in the valuation study. This phenomenon is most evident in stated preference studies that sample from a particular geographic region, e.g., a state and ask participants about a particular, compartmentalized landscape. Studies such as these censor willingness to pay observations that are outside of that state and thus only estimate aggregate values conditional on the sampled population. These studies also compartmentalize the landscape by assuming that people's preferences or landscape structure and function can be meaningfully geographically isolated. By dividing the aggregate welfare measure by the count of the modeled population, we control for the arbitrary influence of censoring WTP of non-sampled populations that have positive WTP. We also control for the effect of the sampled population by including this as an explanatory variable. Analogously, we normalize the aggregate welfare measures by the number of wetland acres valued in the study and include this acreage count as an explanatory variable.

The PLWLS Regression

The majority of the technical aspects of our discussion of the novel estimator is confined to the online appendix[1]. The purpose of this section is to lay out an intuitive understanding of how the parametric locally weighted least squares (PLWLS) approach works; we assume the reader is familiar with the OLS estimator. The appendix describes a two-step process, the first is calibration where an unconventional type of parameter is estimated, and the second step is the use of the estimated parameter to develop weights for a weighted least squares regression.

The motivation behind developing the PLWLS estimator and going beyond OLS is that we wish to estimate a regression that is targeted towards each site (i.e., a site of interest or centered site) for which we desire an estimate of value. The motivation for selecting a potentially unique model for each site comes from a desire for forecast efficiency, or low prediction error variance. The basic concept behind the PLWLS regressor is that certain variables, referred to as *correspondence attributes*, are indicative of how *close* observations are to each other. Two observations that are *close* to each other provide more information about each other than two observations that are not. We make use of this correspondence information by applying lower weights to data that are less *close* to the site of interest. The idea of correspondence has arisen in a number of ecosystem service MA publications, but we are unaware of any attempts to formalize a systematic procedure for estimating or implementing correspondence.

An immediate challenge that arises in implementing this idea is that we do not know how to combine *correspondence attributes* with very different units. Thus we estimate parameters that serve as a normalization that allows us to combine dissimilar *correspondence attributes*. For example if we suspect that geographic distance between two sites and the difference between two sites in average income of the local population are both important indicators of correspondence, lacking a quantitative model, we have only our best judgment to guide us in deciding the relative importance of distance and difference in income. With PLWLS we are able to estimate parameters, referred to as *correspondence parameters*, that allow us to add up the *correspondence distance* between two sites as measured by both geographic distance and the difference in income. An important challenge is that there is a cost to applying lower weights to certain observations, as we would be losing useful information if we incorrectly reduce the weight of a particular observation in the model for a site of interest. Accordingly, we allow *correspondence parameters* to equal zero (which results in the same OLS estimator in each regression) or any positive, real number. We also analyze the performance of the estimator in comparison to OLS in an artificial forecast simulation, described in the next section.

The main divergence of our model from a conventional OLS regression is that we have a regression for each observation; these regressions are distinguished by potentially unique weights used during estimation. Equation (2) contains the basic formula for estimating the weight applied to observation j in the regression tailored to site i. In this equation we have the sum of H terms in a negative exponential function.

$$\widetilde{\boldsymbol{\Omega}}_{i\,jj}^{-1} = e^{-\sum_{h=1}^{H}|a_{hi}-a_{hj}|\tilde{\delta}_h} = e^{-(|a_{1i}-a_{1j}|\tilde{\delta}_1+\cdots+|a_{Hi}-a_{Hj}|\tilde{\delta}_H)} \qquad (2)$$

The variable a_{hi} is a correspondence attribute where the subscript h indicates which attribute (e.g., income or geographic location) and the subscript i indicates the observation associated with the attribute. Thus $|a_{hi} - a_{hj}|$ is an unweighted measure of correspondence distance, which is weighted by the estimated correspondence parameter, $\tilde{\delta}_h$, and the sum of these terms is the argument of the negative exponential function. The resulting value, $\widetilde{\boldsymbol{\Omega}}_{i\,jj}^{-1}$, is the j[th] diagonal element of the diagonal, positive definite regression weight matrix, $\boldsymbol{\Omega}_i^{-1}$. The desired regression parameter vector, $\tilde{B}_i$, is calculated according to the WLS formula in equation (3), where the tilde above a variable indicates that we have estimated that variable via L-WLS.

$$\tilde{B}_i = (\boldsymbol{X}'\widetilde{\boldsymbol{\Omega}}_i^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\widetilde{\boldsymbol{\Omega}}_i^{-1}Y \qquad (3)$$

Where X is an nxk matrix of explanatory variables including an intercept term, Y is a column vector of observations of WTP of length n, and $\tilde{B}_i$ is an n-element column vector of parameter estimates. The subscript i indicates that the matrix $\widetilde{\boldsymbol{\Omega}}_i^{-1}$ has been estimated for observation i; the weight matrix is the source of information that leads to a potentially unique parameter estimate for site i.

[1] http://tinyurl.com/PLWLSappendix

RESULTS

OLS Results

The results of our OLS regression can be found below in Table 1.

Table 1: OLS MA Regression Model Results

Variable	B OLS	s.e.	t-stat
Intercept	24.147	19.863	1.216
Study/Site Variables			
Acres valued	-0.583*	0.306	-1.904
Population	-0.41***	0.148	-2.761
Methodological Variables			
Revealed preference	-4.401*	2.319	-1.898
Stated preference	0.864	1.305	0.662
Joint valuation (1,2,3)	-5.977***	1.435	-4.164
Water quality	5.969**	2.864	2.084
Flood protection	5.438*	2.861	1.9
Total value	6.534**	2.836	2.304
Recreation, general	10.667***	2.627	4.061
Habitat	6.564**	3.093	2.122
Recreation, fishing	5.287**	2.178	2.428
Recreation, hunting	5.011*	2.628	1.907
Interaction_use	-1.211**	0.562	-2.153
Interaction_passive	-0.705	0.56	-1.258
Context Variables			
GDP(state)	-2.681	2.076	-1.291
Coastal (1/0)	1.214**	0.566	2.144
GDP*local_pop	0.038	0.029	1.309
Local_pop:local_wet	2.394**	1.124	2.13
R^2 = 0.81			
*Significance levels: ***<0.01, **<0.05, *<0.1*			

In Table 1, we report Huber-White or heteroskedasticity robust standard errors and the t-statistic for the null hypothesis that each parameter is equal to zero. The parameters estimated for both water quality and flood control are significant at the 10% level or better. While the water quality parameter is significant at the 5% level, the flood control parameter just barely misses that mark, suggesting that the parameter estimates are reasonably precise. The parameter estimate for the number of acres valued is negative and significant at better than the 10% level with robust standard errors. The parameter for population is similar in magnitude to the parameter for acres, but the p-value is less than 1%. Because the dependent variable is normalized by acres and population, the interpretation is that a 1% increase in one of these variables leads to a ~0.5% decrease in the WTP/acre/person, which suggests diminishing returns to expanding the scope of valuation or the scope of the population for aggregation. The results of our OLS forecasts are discussed with the results of the PLWLS estimator in the next sub-section as well as in the discussion section.

PLWLS Results

As the calibration of *correspondence parameters* is the first step in the PLWLS estimator, we present these values first. The results of our Locally Weighted Least squares calibration can be found in Table 2.

Table 2: PLWLS Estimated Correspondence Parameters

Correspondence Attribute	Estimated Correspondence Parameter
Flood control	1.365992
GDP	1.276706
Water quality	0.760224
Coastal	0.218638
Distance 10^3km	0.194442
Population	0.133756
Acres	0

All *correspondence attributes* were standardized by their sample means and variances prior to calibration, with the exception of distance which was converted from meters to 1000's of kilometers. In Table 2 we have removed the transformation for correspondence attributes other than distance to facilitate a comparison. GDP, and distance between sites are important non-methodological determinants of correspondence among studies in our sample. Additionally, studies that value similar services and which aggregate over a similar population are also important determinants of correspondence. The PLWLS calibration step found that the number of acres valued by a study were not determinants of correspondence, a notable result that might be reassessed in future studies. The moderate values for these parameters and the moderate weights (typically between 0.01 and .9) that are consequently applied in each regression suggest that the algorithm is moderately down-weighting observations with poor correspondence, but retaining ample information for reasonably robust estimation.

The next step in the PLWLS procedure after identifying correspondence parameters is to use those parameters to calculate an nxn matrix of regression weights for each out-of-sample study/site of interest. These regression weights are used to calculate a regression for each site and also can be used to rank observations according to their relative information content. For any given centered site, observations that are weighted more heavily are assumed to have greater correspondence and therefore be more informative about the centered site.

We present the valuation forecasts of a stated preference study of flood control and water quality benefits for each of our NWR policy sites in Table 3. The table contains the median and mean of the dependent variable after reversing the log transformation and the annual values are scaled to 2010 US dollars per thousand acres per thousand people. The population is specified as the mean value in our dataset, about 3.5 million people. The median values are the point estimates of WTP forecasted by OLS and PLWLS estimators. For each estimator we include a forecast of both the median and mean values of willingness to pay. The magnitude of the difference between the median and the mean is inversely proportional to the precision of each model, and the consistently smaller gap between mean and median for the PLWLS model implies that this model is consistently more precise than the OLS model.

Table 3: OLS and PLWLS Forecasts for 4 NWRs, Annual dollars per 1000 Acres per 1000 People per Year

2010 US dollars per 1000 acres per 1000 people per year		OLS		PLWLS	
Site	Service	Med	Mean	Med	Mean
Arrowwood NWR	WQ	$170	$520	$370	**$450**
	FC	$100	$310	$110	**$130**
Blackwater NWR	WQ	$720	$2,210	$290	**$330**
	FC	$420	$1,300	$490	**$550**
Okefenokee NWR	WQ	$160	$480	$80	**$90**
	FC	$90	$280	$180	**$210**
Sevilleta & Bosque NWRs	WQ	$320	$970	$810	**$980**
	FC	$190	$570	$210	**$260**

Also of note in Table 3, many of the PLWLS dependent variable estimates, both mean and median, fall between the mean and median of the OLS estimates, implying that the PLWLS results do not suffer from extensive bias relative to the results from the unbiased OLS estimator. We highlight the mean value single service revealed preference study results as we consider these to be the best estimates (i.e., for forecasting the value of ecosystem services when estimated by a primary valuation study) available from our MA regression methodology. Under our suggested interpretation of the forecasts of the dependent variable, best estimates include the judgment that a stated preference study of a population with about 3.5 million people would be the best choice to estimate the total economic value of each service.

The mean of our PLWLS estimator is what we consider to be our best forecast, conditional on the population count used during estimation. An important contribution of our paper is the modeling and associated acknowledgement that the population over which welfare estimates are aggregated in primary valuation studies is essentially always a choice of the original analyst and dictated neither by the model chosen by the analyst nor the context of the site associated with the ecosystem services being valued. Studies that utilize an empirical approach to restricting the population over which benefits are aggregated are surprisingly rare; Sutherland and Walsh's (1985) study is the only domestic exception we encountered in our literature search. Clearly an important next step is to develop a more formal empirical means for choosing the population over which benefits are aggregated.

Essentially, we follow the existing MA literature in agreeing that methodological covariates can be difficult to assign for forecasting ecosystem service values when one is simply interested in knowing how an acre of wetlands impacts the welfare of society. We expect that the estimation of a single number that is free of methodological underpinnings is not the goal of MABT, not anytime soon. Rather, we reiterate that the best interpretation of the dependent variable obtained from a MA regression model is a simulated study result. As ecosystem service valuation studies are typically conducted in the context of answering a research or policy question, one must choose a specific valuation methodology to simulate a study result. If interest lies in a value that is an average of results obtained from a variety of methodologies, we recommend simulating each valuation methodology by coding methodological dummy variables to 1 or 0, and taking the average (perhaps with unequal weights) of the final forecasts. We specifically caution against coding dummy variables for binary concepts as fractions(e.g., coding them at their sample means); if this caution is not followed, the resulting forecasts will be non-linear functions of multiple valuation methodologies that has no clear interpretation and no means for empirical validation.

The results of the regressions using our dataset produces numbers with a magnitude most comparable to the values estimated by Woodward and Wui (2001). In comparison to the valuation forecast results obtained from the median of the Ghermandi et al. (2010) MA model, our OLS values tend to be about 50 times higher. The mean values simulated for the Ghermandi et al. (2010) study are substantially closer to the PLWLS mean values obtained with our dataset, though substantial differences exist. Because we do not have access to information regarding even the average population in the datasets associated with the 3 previously published MAs, we cannot confidently say that the majority of the variation in benefits is not due to a difference in the population that was used to make the forecast.

CONCLUSIONS

The process of learning about the behavior of the PLWLS estimator in our sample has led us to conclude that the most telling way to understand the estimator is in the context of penalties. Starting from an OLS regression where no observation is penalized, when we consider a local regression for site i, observations that are distant (in terms of correspondence) from the central observation are penalized. The penalties are unity less the weights determined by the parameterized exponential equation described in the appendix. The process of calibration works by fitting a parametric variance equation to sample variances and efficient forecasts are made based on patterns identified in the calibration process. The results of our PLWLS estimator seem promising, yet the impact of clustered observations and a small sample size on the performance of the model along with a reliable means for estimating regression coefficient variability due to resampling are still important unanswered questions.

The appropriate use of welfare measures is dictated by the questions one wishes to answer. The relative lack of attention to the difference between mean and median values of the dependent variable (e.g., Woodward and Wui 2001; Brander, Florax, and Vermaat 2006; Ghermandi et al. 2010; Brander, Brouwer, and Wagtendonk 2013) is problematic as these two measures of central tendency are appropriate for answering different questions. Median values may be useful for predicting the outcome of a vote for which an outcome requires a simple majority. Benefit-cost analysis and in general estimation of economic benefits on the other hand typically require that one uses the value of the mean of the dependent variable. For models estimated with the dependent variable in log form, estimation of the median value of the dependent variable is easy, but this ease comes at the cost of potentially using the wrong value (biased downwards) which for benefit-cost analysis will systematically lead to biased decisions that lead to a reduction in social welfare as measured by aggregate benefits due to inadequate protection and restoration of wetlands that provide important services.

Future work will likely increase the efficiency of benefit transfers. Efforts to increase efficiency can focus on expanding the MA dataset through several avenues: inclusion of non-domestic studies and appropriate controls, increasing the sample size by adding additional services to the analysis, and increasing the sample size by more carefully reviewing existing studies and relevant technical reports for missing information that prohibited their use in the current analysis. Validation and updating of the model through future primary valuation studies is also an important aspect of increasing the efficiency of MABT. Work is ongoing to identify common domestic wetland sites that have poor correspondence with the existing MA dataset, as primary valuation studies of these sites are likely to greatly enhance our ability to produce efficient forecasts. Increasing the efficiency of MABT models for wetland sites that are both common and understudied is a particularly promising avenue for future primary valuation studies.

REFERENCES

Bishop, R.C. 2003. “Where to from Here?” In *A Primer on Nonmarket Valuation*, edited by P.A. Champ, Kevin J. Boyle, and T.C. Brown, 3:537–566. The Economics of Non-Market Goods and Resources. Norwell, MA, U.S.A.: Kluwer Academic Publishers.

Brander, Luke, Roy Brouwer, and Alfred Wagtendonk. 2013. “Economic Valuation of Regulating Services Provided by Wetlands in Agricultural Landscapes: A Meta-analysis.” *Ecological Engineering* in press.

Brander, Luke, Raymond Florax, and Jan Vermaat. 2006. “The Empirics of Wetland Valuation: A Comprehensive Summary and a Meta-analysis of the Literature.” *Environmental and Resource Economics* 33: 223–250.

Champ, P.A., Kevin J. Boyle, and T.C. Brown. 2003. *A Primer on Nonmarket Valuation*. Vol. 3. Springer Netherlands.

Ghermandi, A., J. Van den Bergh, Luke Brander, H.L.F. de Groot, and P. Nunes. 2010. “Values of Natural and Human-made Wetlands: A Meta-analysis.” *Water Resources Research* 46 (12): 1–12.

Woodward, RT, and YS Wui. 2001. “The Economic Value of Wetland Services: a Meta-analysis.” *Ecological Economics* 37: 257–270.