University of Rhode Island

DigitalCommons@URI

Open Access Dissertations

2018

ESSAYS ON ECONOMICS OF ECOSYSTEM SERVICES FROM DAM REMOVAL TO IMPROVE DECISION MAKING

Ben Blachly

University of Rhode Island, benblachly@gmail.com

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Recommended Citation

Blachly, Ben, "ESSAYS ON ECONOMICS OF ECOSYSTEM SERVICES FROM DAM REMOVAL TO IMPROVE DECISION MAKING" (2018). *Open Access Dissertations*. Paper 1191.
https://digitalcommons.uri.edu/oa_diss/1191

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

ESSAYS ON ECONOMICS OF ECOSYSTEM SERVICES
FROM DAM REMOVAL TO IMPROVE
DECISION MAKING
BY
BEN BLACHLY

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
ENVIRONMENTAL AND NATURAL RESOURCE ECONOMICS

UNIVERSITY OF RHODE ISLAND

2020

DOCTOR OF PHILOSOPHY DISSERTATION
OF
BEN BLACHLY

APPROVED:

Dissertation Committee:

Major Professor Emi Uchida

Todd Guilfoos

Arthur Gold

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2020

ABSTRACT

Many dams throughout New England are approaching the end of their engineered lifespan. Individual dam owners and governments at various levels find themselves navigating decisions to repair, remove, or modify the aging infrastructure. These decisions have implications for the ecosystem services that depend on the presence or absence of dams. By coordinating the ecosystem service tradeoffs at large scales we can more efficiently utilize the productive capacity of river systems. Implementing a large-scale coordinated approach, however, requires understanding stakeholder preferences at different scales, and the willingness of decision makers to coordinate (or not). In this dissertation I address issues surrounding a coordinated approach to ecosystem service provision with the goal of facilitating better decisions. In the first chapter, I administer a choice experiment survey to study how environmental risk preferences interact with scale to determine willingness to pay for ecosystem services from dam removal. Improved understanding of scale-dependent preferences can allow nonmarket valuation estimates for ecosystem services to integrate more smoothly into decisions at a variety of scales. In the second chapter, I explore environmental federalism in dam removal decisions (i.e., whether state goals diverge from larger-scale optimization). I use matching and instrumental variable techniques to model the determinants of dam removal using a large, spatially explicit dataset. Results suggest that states consider border proximity when selecting dams for removal, indicating a need for new incentive structures to realize efficiency gains from coordination. In the final chapter, I study how a production possibilities model can be combined with public preference data from a choice experiment, expressed as

indifference curves, to identify socially preferred ecosystem service outcomes from dam removal. I find that the approach is useful for pinpointing areas of agreement and disagreement between stakeholder groups with varying preferences.

ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Emi Uchida. Your tireless work ethic, uncompromising standards, humility, and your ability to form personal connections are just some of the talents that have made you an amazing mentor. Your spirit for collaboration and excitement for interdisciplinary work are unmatched. You have so many qualities that I will seek to emulate throughout my career. The opportunity to learn from you would have been enough, but you also provided me with so many tangible opportunities. I will always be grateful for everything you have done for me.

I also wish to thank the other members of my committee. Dr. Todd Guilfoos, you are an amazing economist and I am constantly impressed by your work. I am so appreciative of your support and candid advice over the course of my PhD training. The same goes for Dr. Arthur Gold. Your wit and perspective are unparalleled and always appreciated.

In addition, I want to thank all the current and former graduate students I had the pleasure of meeting. This dissertation would not have been possible without the many conversations we shared. I learned something important from everyone, but I would like to thank Vasu Gaur, Jason Walsh, and Mike Weir in particular for their friendship and support.

Thank you to my father, Dan, who fostered my passion for the environment, and to my mother, Andre, who nurtured my critical thinking.

Finally, I am quite certain this dissertation would not have been possible without constant support and patience from my wife, Aimee, and our two boys, Connor and Oliver. To my entire family: thank you. I love you all so much.

PREFACE

I use manuscript format for this dissertation. There are three independent chapters which constitute the entire work. The goal of this dissertation is to apply economic concepts to improve the ecosystem service outcomes associated with dam removal decisions. The first chapter is co-authored with Emi Uchida and Samuel G. Roy. It is being prepared for submission to the *Environmental and Resource Economics*. The second manuscript is co-authored with Emi Uchida. It is being prepared for submission to *Land Economics*. The third manuscript is co-authored with Emi Uchida and Samuel G. Roy. It is being prepared for submission to *Ecology and Society*.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iv
PREFACE.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
THE RELATIONSHIP BETWEEN ENVIRONMENTAL RISK	
AVERSION AND SCALE: IMPLICATIONS FOR	
BENEFIT TRANSFER AND SCOPE.....	
Abstract.....	1
1.1 Introduction.....	2
1.2 Background and Hypotheses.....	3
1.3 Study Context and Survey Details.....	9
1.4 Methodology.....	15
1.5 Results.....	21
1.6 Discussion.....	26
1.7 Conclusion.....	31
References.....	34
ENVIRONMENTAL FEDERALISM IN RIVER RESTORATION.....	37
Abstract.....	52
2.1 Introduction.....	53
2.2 Background.....	54
	56

2.3 Environmental Federalism and Dam Removal Decisions.....	58
2.4 Data.....	60
2.5 Empirical Strategy.....	63
2.6 Results.....	67
2.7 Discussion.....	71
2.8 Conclusion.....	73
References.....	76

INTEGRATING PUBLIC PREFERENCES WITH BIOPHYSICAL

PRODUCTION POSSIBILITIES: AN APPLICATION TO ECOSYSTEM SERVICES FROM DAM REMOVAL.....	89
Abstract.....	90
3.1 Introduction.....	91
3.2 Concepts and Motivation.....	95
3.3 Study Context.....	98
3.4 Methods.....	100
3.5 Results and Discussion.....	105
3.6 Conclusion.....	112
References.....	115

LIST OF TABLES

TABLE	PAGE
Table 1.1 Attribute levels for both scale treatments.....	45
Table 1.2 Demographic distributions by scale treatment.....	46
Table 1.3 Utility specifications.....	47
Table 1.4 Mixed logit results.....	48
Table 1.5 Transfer errors.....	49
Table 1.6 Mixed logit results for uncertainty subsample.....	50
Table 1.7 Mixed logit results for consequentiality subsample.....	51
Table 2.1 Result of t-tests for pre-matching difference in means.....	85
Table 2.2 Result of t-tests for post-matching difference in means.....	86
Table 2.3 Marginal effects from logistic regression.....	87
Table 2.4 Instrumental variable results.....	88
Table 3.1 Regression results.....	127
Table 3.2 Descriptive statistics by class.....	128
Table 3.3 Two-dimensional utility maximizing combinations by class.....	129
Table 3.4 Four-dimensional utility maximizing combinations by class.....	130
Table 3.5 Welfare estimates from two- and four- way optimizations.....	131

LIST OF FIGURES

FIGURE	PAGE
Figure 1.1 Geographic scale treatment.....	42
Figure 1.2 Sample choice set from the subwatershed treatment.....	43
Figure 1.3 External scope test results.....	44
Figure 2.1 Dam removal frequency by year.....	81
Figure 2.2 Number of existing and removed dams by state.....	82
Figure 2.3 Geospatial distribution of treatment and control groups.....	83
Figure 2.4 Comparison of densities before and after matching.....	84
Figure 3.1 The Penobscot Watershed.....	119
Figure 3.2 Sample choice set.....	120
Figure 3.3 Two-dimensional PPFs.....	121
Figure 3.4 Four-dimensional optimization results plotted in two dimensions..	122
Figure 3.5 Marginal WTP.....	123
Figure 3.6 PPF combined with MNL results.....	124
Figure 3.7 Hydro/lake vs salmon PPFs with indifference curves by class.....	125
Figure 3.8 Hydro vs herring/salmon PPFs with indifference curves by class...	126

The relationship between environmental risk aversion and scale:
Implications for benefit transfer and scope

(To be submitted to Environmental and Resource Economics)

Ben Blachly¹, Emi Uchida, Samuel G. Roy

¹ Ben Blachly is corresponding author and is a PhD candidate in the Department of Environmental and Natural Resource Economics, University of Rhode Island (ben_blachly@uri.edu). Emi Uchida is a Professor in the Department of Environmental and Natural Resource Economics, University of Rhode Island. Samuel G. Roy is an Assistant Research Professor at the Senator George J. Mitchell Center for Sustainability Solutions, University of Maine. Support for this project is provided by the National Science Foundation's Research Infrastructure Improvement NSF #IAA-1539071. We thank participants at the Association of Environmental and Resource Economists 2019 Summer Conference for useful comments.

ABSTRACT

In stated preference studies, willingness to pay estimates sometimes fails to increase with the quantity of an environmental good. This theoretical inconsistency leads to errors when performing benefit transfer and has been used to question the validity of nonmarket valuation. This study tests whether this inconsistency can be partially explained by how individual risk preferences interact with geographic and quantitative scales. To address this question, we administer a split-sample choice experiment with a scale treatment, and model the data using different assumptions about risk aversion. We find that respondents exhibit a higher degree of risk aversion at larger geographic and quantitative scales of the nonmarket good, and a model incorporating risk aversion eliminates between-scale benefit transfer errors for a subset of respondents. In addition, we subject our choice experiment to a novel between-sample scope test, showing that the interaction of risk aversion with scale is a significant driver of scope response. Our results make the case for including a risk preference elicitation among standard debriefing questions in stated preference surveys.

1. Introduction

In stated preference studies for nonmarket goods, willingness to pay sometimes lacks sensitivity to the quantity of the good on offer (Desvousges et al., 2012). Known as failure to demonstrate scope in the contingent valuation literature, a number of studies were unable to demonstrate increasing willingness to pay for increasing quantities of an environmental good (e.g., Boyle et al., 1994; Fischhoff et al., 1993; Kahneman and Knetsch, 1992). In a discrete choice experiment setting, marginal willingness to pay may depend on the scale at which it is elicited (Bueren and Bennett, 2004; Rolfe and Windle, 2008). If marginal values derived at a relatively small (large) scale are relatively large (small), estimates of total value could be decreasing with the magnitude of environmental change. Scale dependency of marginal values in choice experiments are virtually unstudied relative to scope response in contingent valuation, despite their relatedness. We explore scale dependency in the context of uncertainty, an increasingly common component of choice experiments.

Uncertainty is inherent in environmental decision-making (Pindyck, 2006). Respondents to stated preference surveys face uncertainties from several sources surrounding the ultimate delivery of the outcome being valued (e.g., environmental, modeling, and political uncertainty). A growing number of nonmarket valuation studies show that explicit inclusion of uncertainty in a valuation task alters preferences (e.g., Lundhede et al., 2015; Makriyannis et al., 2018; Roberts et al., 2008). Still, our understanding of how people respond to environmental uncertainty is lacking compared to well-established domains like financial uncertainty. Of relevance to our

investigation of scale dependency is that people become more averse to financial risk as the scale of payoffs becomes larger (Binswanger, 1980; Bosch-Domènech and Silvestre, 1999; Holt and Laury, 2002).

Our paper studies the relationship between environmental risk aversion and scale, and explores implications for stated preference nonmarket valuation studies. Using choices between incentivized lotteries in a lab, Holt and Laury (2002) found subjects made fewer risky choices when the monetary stakes were scaled up. Their finding confirmed earlier evidence that people exhibit a higher degree of risk aversion when the financial payoffs are larger (Binswanger, 1980; Bosch-Domènech and Silvestre, 1999). For stated preference studies valuing uncertain environmental outcomes, greater discounting of risk at larger scales could cause the related issues of scope failure in contingent valuation and scale dependency in choice experiments. However, it is unknown whether the same relationship holds in an environmental context.

Our paper contributes to the growing literature on preferences for environmental risk and how they compare to preferences for financial risk. Research using both choice experiment data (Bartczak et al., 2016; Faccioli et al., 2019; Holzer and McConnell, 2017) and adapted lottery menus (Bartczak et al., 2015; Riddell, 2012) supports the general notion that people are averse to environmental risk. Parallels between financial and environmental risk, however, are not straightforward. For example, prospect theory has shown that people exhibit probabilistic aversion to financial risks by underweighting low probabilities and overweighting high probabilities (Kahneman and Tversky, 1979). In an environmental context, some

researchers have found a similarly-shaped probability weighting function (Bartczak et al., 2015; Riddel, 2012), while others have found the inverse (Roberts et al., 2008).

We focus on whether the established positive relationship between financial risk aversion and scale carries over to an environmental setting, since it has important implications for stated preference research. We detail the implications below.

We contribute to the expansive scope literature by suggesting the interaction of risk aversion with scale as a new explanation and subjecting a discrete choice experiment to a rare between-sample scope test. Explanations for scope failure have been explored extensively in the contingent valuation literature (see Lopes and Kipperberg, n.d. for a complete review). Quickly diminishing marginal utility (Rollins and Lyke, 1998) and the purchasing of moral satisfaction for a flat fee (Kahneman and Knetsch, 1992) are among the most widely recognized explanations. To our knowledge, only one study in the literature relates scope response to scale and uncertainty, but it does not allow for risk aversion. Powe and Bateman (2004) discovered that survey respondents perceive a lower provision probability when an environmental program is proposed at a larger geographic scale, which attenuates scope response. While scope failure has received much scrutiny in contingent valuation, it has been largely ignored in the discrete choice literature. In fact, we know of just one study subjecting a choice experiment to a between-sample (external) test for scope response (Andersson et al., 2016), which is considered the standard in contingent valuation (Arrow et al., 1993).

Finally, we contribute to the benefit transfer literature by testing whether models that incorporate risk aversion can reduce between-scale transfer errors. Choice

experiments are attractive for benefit transfer, since their hedonic structure allows extracting marginal values for individual attributes (Rolfe et al., 2015). However, if marginal values are scale dependent, some sort of calibration is required to reduce the likelihood of transfer errors (Bueren and Bennett, 2004; Rolfe et al., 2013). Because risk aversion may contribute to scale dependency, we test the ability of a model capturing risk averting behavior to reduce between-scale transfer errors compared to a model assuming risk neutrality.

To explore these topics, we administer a split sample choice experiment with a geographic scale treatment. We create two versions of the experiment, one framed at a large watershed scale in Maine and the other at a smaller, embedded subwatershed scale. We elicit preferences for expanding access to spawning habitat for endangered Atlantic salmon by removing dams at the expense of hydropower capacity. We vary the likelihood that salmon will reach the numbers expected by scientists, reflecting uncertainty surrounding recovery efforts. We simulate feasible levels for the attributes at both geographic scales using ecological models (Roy et al., 2018). Given that the subwatershed is embedded within the watershed, we have a confounding of scope (the amount of an attribute, or the magnitude of a change) with scale (the geographic range over which the attribute or change extends). Henceforth we will use scope to describe the theoretical property of willingness to pay that increases with the quantity of an environmental good, and scale to indicate both the geographic extent and quantitative range of an attribute (i.e., scope is satisfied when willingness to pay increases with scale).

We model the choice experiment data at both scales using different assumptions about risk aversion and test whether environmental risk aversion increases with scale, and the implications for understanding the results of a scope test and transferring benefits between scales. Our baseline model is a linear expected utility specification, which implies risk neutrality. Then we allow the utility function to exhibit concavity by specifying a constant relative risk aversion functional form and estimating a risk aversion coefficient. Our final model allows for heterogeneous risk attitudes self-reported by respondents in the questionnaire. We use the three model results from both scale treatments to obtain welfare estimates for restoring Atlantic salmon to the subwatershed and watershed. To answer our research questions, we make comparisons between the relevant model parameters and welfare estimates at both scales.

We find that respondents exhibit a higher degree of environmental risk aversion at the larger scale. Using our nonlinear model, we estimate risk aversion coefficients of 0.694 and 0.923 at the subwatershed and watershed scale, respectively. Additionally, our model that incorporates self-reported risk attitudes indicates a relatively small and statistically insignificant discount associated with risk aversion at the subwatershed scale. At the watershed scale, however, the risk aversion discount grows in magnitude and becomes statistically significant. We take this as strong evidence that risk aversion and scale have the same relationship in the environmental domain that has been established in the financial domain (Binswanger, 1980; Bosch-Domènech and Silvestre, 1999; Holt and Laury, 2002).

Turning to scope, we find that the magnitude of scope response varies with the degree of risk aversion. Respondents who self-identify as highly risk averse fail an external scope test (i.e., willingness to pay decreases with scale), while those identifying as more tolerant to risks exhibit a strong between-sample scope response (i.e., willingness to pay strictly increases with scale). Andersson et al. (2016) administered a similar split-sample scope test in the context of valuing a statistical life, using a choice experiment to understand preferences for varying levels of mortality risk. They report evidence of scope failure. Our results suggest risk preference interacting with scale may be contributing to the lack of scope in their study, which is firmly in a risky context. In general, we offer the interaction of risk preferences and scale as a novel consideration when using scope to evaluate the validity of a study.

Finally, we find a model that incorporates risk aversion eliminates between-scale benefit transfer errors for a subsample of our respondents. A risk neutral model produces scale-dependent implied prices, which result in significant between-scale transfer errors. Welfare estimates for the risky attribute, Atlantic salmon, are about twice as large using the subwatershed-derived parameters compared to the watershed-derived parameters. By utilizing a choice model that captures risk averting behavior, we eliminate between-scale transfer errors for a subsample of respondents who perceive uncertainty beyond what is explicit in the survey. Transfer errors are reduced, but not eliminated, for another subsample who view the survey as consequential. Utilizing our full sample, the transfer errors increase when we apply the risk aversion model. These mixed results suggest that people may have different preferences for risk

that arises from different sources. Further research is needed to operationalize risk preferences as a viable approach for improving benefit transfer.

2. Background and hypotheses

2.1 Scale dependency and uncertainty

First, we define scale in the context of this paper. We use scale to refer to both the geographic and quantitative range of an attribute, which are linked by an ecological production function. For example, carrying capacity is an increasing function of habitat area. While it is possible to separate geographic and quantitative scale, for example by altering population density, we ensure inseparability by carefully defining each attribute.

When environmental economists have manipulated scale in a split sample choice experiment, they find lower (higher) implied prices larger (smaller) scales. We refer to this inverse relationship as scale dependency. Bueren and Bennett (2004) conducted a choice experiment to value land and water degradation in Australia, testing the effect of regional versus national framing. Holding population constant, they find the magnitude of implied prices is larger at the smaller (regional) scale. Rolfe et al. (2013) find marginal values for conservation are significantly lower for larger geographic frames in the context of both the Great Barrier Reef and an Australian river catchment. In contrast, Rolfe and Windle (2008) report no difference between marginal values elicited at regional and state scales. However, they presented

attributes as percentages (i.e., the percentage of a region's soil that will be allowed to degrade). Converting percentages to the appropriate numeric quantities results in scale dependency. Finally, Schaafsma and Brouwer (2013) also test the effect of a geographic scale treatment, although the goal and design of their study differs significantly from ours. Nonetheless, they report lower point estimates for implied prices associated with broader geographic scales.

We build on this research by testing for scale dependency of an explicitly uncertain ecosystem service attribute. The outcomes from environmental programs described in stated preference surveys are subject to several real and perceived uncertainties from a respondent's perspective: environmental uncertainty, political uncertainty, and modeling uncertainty. While some unknowns are truly ambiguous, others may be approximated by probabilities. For example, we may know that a particular restoration strategy has roughly a 50% success rate. Split-sample studies show that failing to include relevant risk information in a valuation task alters preferences, leading to incorrect welfare estimates (e.g., Lundhede et al., 2015; Makriyannis et al., 2018; Roberts et al., 2008; Rolfe and Windle, 2015). At the same time, there is an apparent relationship between uncertainty and scale.

Limited research indicates a direct relationship between environmental outcome uncertainty and scale. In a contingent valuation survey, Powe and Bateman (2004) discovered that respondents found restoration outcomes framed at larger scales less realistic. We contribute to the nascent literature investigating environmental uncertainty and scale by testing the following initial hypothesis.

H1: Marginal willingness to pay for an explicitly uncertain attribute is lower (higher) at a larger (smaller) scale.

2.2 Environmental risk aversion and scale

The robust evidence of scale dependency reported above may result from respondents discounting perceived or explicit outcome uncertainty to a higher degree at larger scales. Behavioral evidence exists that people become more risk averse when the scale of financial payoff is larger. In a classic set of laboratory experiments, Holt and Laury (2002) elicited subjects' baseline risk preferences through their preferences for incentivized lotteries. Next, they repeated the experiment but with the lottery payoffs scaled up by a factor of 90. At the higher payoff levels, subjects tended to prefer safer lotteries (i.e., they were more risk averse). Their result confirmed earlier findings from the lab (Bosch-Domènech and Silvestre, 1999) and in the field with rural farmers in India (Binswanger, 1980). These findings, though, are all in the context of financial risk.

Generally, research on risk preferences in an environmental context has found that people are not neutral to environmental risk. For example, people exhibit aversion to risky conservation outcomes (Faccioli et al., 2019) and random catch in recreational angling (Holzer and McConnell, 2017). Related research investigates whether preferences for financial lotteries are predictive of risky environmental choices. Grijalva et al. (2011) found financial risk aversion to be associated with support for a safe minimum standards approach to resource management, and Bartczak et al. (2016)

provide evidence of a latent variable driving both financial lottery and risky endangered species conservation decisions. Finally, a few papers directly compare financial lottery choices to choices in lotteries adapted to an environmental context, noting similar risk preferences across the two domains (Bartczak et al., 2015; Riddel, 2012).

Our paper contributes to the environmental risk preference literature by testing for a relationship between risk aversion and scale in an environmental context. In a choice experiment eliciting bird conservation preferences, Lundhede et al. (2015) found additional disutility from outcome uncertainty at the highest levels of conservation. We build on this initial evidence of a positive environmental risk aversion-scale relationship by conducting a split-sample test of the following hypothesis.

H2: Choice experiment respondents exhibit a higher degree of risk aversion at a larger scale.

In the next two subsections we describe implications of scale-dependent marginal willingness to pay, and what we might learn by considering the environmental risk aversion-scale relationship.

2.3 Implications for benefit transfer

Scale-dependent implied prices can be a source of benefit transfer errors. Benefit transfer is the process of taking valuation estimates derived from a primary study or studies and applying them to a target site for which primary valuation data is

unavailable (Johnston et al., 2015). There has been a general trend toward using meta-analysis-type benefit transfer functions, but empirical evidence for the reliability of this approach is mixed (Johnston and Rosenberger, 2010). Extracting a marginal value for an individual attribute from a primary study and applying it to a target site remains an attractive option (Rolfe et al., 2015). However, scale dependency implies that a scale mismatch between policy and target sites will result in transfer errors. In fact, some of the split sample research described in section 2.1 focuses on calibration techniques for the express purpose of reducing between-scale transfer errors (Bueren and Bennett, 2004; Rolfe et al., 2013). But these calibration factors are formed on a case study basis, and no systematic approach exists (Rolfe et al., 2015).

If scale dependency results, at least in part, from respondents' aversion to environmental risk (see Hypothesis 2) then modeling risk aversion may reduce scale dependency. A reduction in scale dependency (i.e., convergence of implied prices derives at different scales) would result in reduced between-scale transfer errors. We state this as a hypothesis below.

H3: Modeling respondents' risk aversion reduces between-scale transfer errors.

2.4 Implications for scope response

Demand theory dictates that respondents should exhibit nondecreasing willingness to pay for increasing amounts of an environmental good or service. In the stated preference literature, demonstrating consistency with this property is known as

demonstrating scope. The National Oceanic and Atmospheric Administration Blue Ribbon Panel identified scope as an important validity test for contingent valuation studies (Arrow et al., 1993), and it is now considered a desirable property for stated preference studies more generally (Johnston et al., 2017).

Countless contingent valuation studies have tested for scope, with mixed results (Desvousges et al., 2012). However, subsequent research has shown that failure of a statistical scope test does not necessarily negate the validity of a study. Indeed there are several theoretically and behaviorally consistent explanations for a muted or nonexistent scope response (see Lopes and Kipperberg, n.d. for a complete review). Among the most widely recognized are quickly diminishing marginal utility (Rollins and Lyke, 1998) and the purchasing of moral satisfaction for a flat fee, deemed the ‘warm glow effect’ (Kahneman and Knetsch, 1992). Of particular relevance to our study, Powe and Bateman (2004) found that conditioning on perceived outcome certainty improved scope test results.

In contrast to contingent valuation, scope has received scant attention in the choice experiment literature. One possible reason is that choice experiments exhibit internal, or within-sample scope response by design. A positive (negative) statistically significant coefficient for an attribute that provides utility (disutility) indicates the presence of internal scope. However, the inverse relationship between marginal values and scale, what we call scale dependency, could imply external scope failure while satisfying the internal measure. Demonstrating external, or between-sample, scope has typically proven more difficult and is considered a more rigorous test (Arrow et al.,

1993). We are aware of just one study that submits a choice experiment to an external scope test, and it found evidence of scope failure (Andersson et al., 2016).

We contribute to the vast scope literature by testing a choice experiment for external scope, and offering risk aversion as a novel contributor to scope response. If respondents exhibit a higher degree of risk aversion at a larger scale (see Hypothesis 2), external scope response is attenuated compared to a situation where respondents weigh risk consistently across scales. This comprises our fourth and final hypothesis.

H4: Risk aversion, and particularly its interaction with scale, is a determinant of external scope response.

3. Study context and survey details

We place our study in the context of valuing ecosystem services from dam removal. The mix of ecosystem services provided by a river are highly dependent on the presence or absence of dams in the river system. Dams were built in America mainly to harness the power of rivers to support industry and provide electricity, but they can also provide flood control and flatwater recreation opportunities, among other benefits. But these benefits come at the expense of different ecosystem services. For example, the negative effect of dams on fish habitat is well-documented (e.g., Poff and Hart, 2002) and there is increasing recognition for the general ecological benefits of freely flowing rivers (Auerbach et al., 2014; Bednarek, 2001). While there has been a marked increase in the number of purposeful dam removals in recent years (American Rivers, 2018), removals remain rare relative to the number of existing dams. Given the

expected growth in river restoration by dam removal, it is important to understand public preferences for the ecosystem service tradeoffs involved.

3.1 Study site and scale treatment

We set our study in the Penobscot watershed located in the state of Maine, USA, which is suitable for a number of reasons. First, the ecosystem services provided by rivers and dams in the Penobscot watershed are conducive to our research questions. The main tradeoff, as identified in focus groups, is between hydropower production and sea-run fish habitat. Hydropower production in Maine is significant. In 2018, nearly one third of net electricity production in the state came from hydropower (United States EIA, 2019). However, the large dams that support hydropower production have consequences for other ecosystem services.

Endangered Atlantic salmon are expected to be a key beneficiary from dam removal in the Penobscot, but their recovery is uncertain. Dams block rivers, prohibiting or limiting the movement of fish throughout the system. Our focus groups identified Atlantic salmon as a particularly high-value species. Atlantic salmon are diadromous, meaning they spawn in freshwater river systems, migrate to the ocean as juveniles, and swim upstream to freshwater habitats to spawn subsequent generations. Dams are considered a major reason for their declining numbers and current listing as an endangered species (Limburg and Waldman, 2009). However, dams are not the only pressure facing Atlantic salmon. Overfishing and climate change have also taken their toll. Perhaps for these reasons, Atlantic salmon have proven a difficult species to

recover (Mills et al., 2013). Though bad for Atlantic salmon, our research design requires this type of outcome uncertainty to utilize choice models that incorporate risk preferences.

In addition, the size of the Penobscot watershed allows us to generate a meaningful scale treatment. The Penobscot is an expansive HUC-04 watershed, originating near the Canadian border and draining about one third of the surface area of Maine (22,196 km²) into the Atlantic Ocean. The watershed is composed of five HUC-08 subwatersheds. We chose the southernmost Lower Penobscot HUC-08 as the second scale to create our treatment. We refer to the broader HUC-04 as simply the ‘watershed’ and the embedded HUC-08 as the ‘subwatershed’ (Figure 1). To ensure that the scenarios are realistic and that the subwatershed levels are nested numerically within the watershed levels, we generate a range of feasible attribute levels for each scale. We describe the methods for generating the levels next.

3.2 Deriving feasible attribute levels at both scales

Following best practices for nonmarket valuation, we design attribute levels at both geographic scales using ecological production functions (e.g., Boyd and Krupnick, 2013). We use the New England Dams Database (<https://ddc-nedams.sr.unh.edu/>) combined with hydropower capacity licensing data from the Federal Energy Regulatory Commission (FERC, 2019), hydrologic data (USGS, 2017), and fish passage and habitat suitability models (Roy et al., 2018) to derive feasible attribute levels at subwatershed and watershed scales. There are 52 and 112

dams in the subwatershed and watershed, respectively. Of these dams, 17 are licensed to generate hydropower, 58 have a lake or pond reservoir, and 62 block historic upstream habitat for Atlantic salmon. We followed the methods in Roy et al. (2018) to measure upper and lower limits of attribute values based on hypothetical scenarios in which all dams in the basin are either kept in current condition or completely removed. While hydropower capacity is simply calculated as the cumulative sum of capacity across all generating dams, Atlantic salmon carrying capacity requires a spatially explicit ecological model that we describe below.

Population estimates for Atlantic salmon are based on available habitat extent, quality, and accessibility data (Roy et al., 2018)

$$S = c \sum_{i \in n_d} \left[h_i \prod_{j \in n_{di}} (p_j) \right] \quad (1)$$

where S is annual Atlantic salmon spawning population capacity (number of fish); n_d is the set of all dams, indexed by i ; n_{di} is the set of all dams downstream from and including dam i , indexed by j ; h_i is the functional habitat accessible above dam i (km^2); p_j is the upstream and downstream product of probability of passage through downriver dam j ; and c is the annual population carrying capacity (number of fish $\text{m}^{-2} \text{a}^{-1}$). We determine the spatial distribution of functional habitat h_i based on physical surveys and historic accounts (Houston et al., 2007; Martin and Apse, 2011), and estimate functional habitat quality based on environmental (Olivero and Anderson, 2008) and hydrologic (USGS, 2017) model data and suitability requirements (Armstrong et al., 2003; NOAA, 2009)

$$h_i = \sum_{r \in g_i} [a_r t_r v_r \vartheta_r \cdot 0.8] \quad (2)$$

where g_i is the set of river reaches immediately above dam i , indexed by r , a_r is the mean annual wetted area within reach r (m^2), t_r is the mean annual temperature factor in reach r , v_r is the mean annual velocity factor in reach r , ϑ_r is a binary value identifying if the river reach r is historically accessible to Atlantic salmon, and 0.8 is included to assume that 20% of available surface area exists along the river banks and is unsuitable for habitat use due to natural river height fluctuations. Locations with seasonal drying or tidal effects are excluded from analysis to reflect species' long-term intolerance of these conditions for spawning and rearing (Armstrong et al., 2003; NOAA, 2009).

3.3 Survey details

The survey included four sections. The first section had questions about recreational usage and quality perceptions of rivers in the watershed or subwatershed. Next was an information section describing how dams impact fish populations and alter recreational opportunities in general, followed by a description of the current situation in the Penobscot. This page differed by scale treatment, showing the current situation in either the watershed or subwatershed. The next section included the choice tasks preceded by a consequentiality statement. The survey concluded with a self-reported measure of risk aversion, questions designed to assess consequentiality and identify protest responses, and a set of standard sociodemographic questions.

Each respondent faced a total of six choice questions. We defined attribute levels based on the feasible range at each of the two scales (Table 1). We used those levels to create a D-efficient design, eliminating the need for a large number of choice sets (Rose and Bliemer, 2009). At each scale we created twelve choice sets broken into two blocks. Each choice set consisted of three alternatives: the status quo and two alternative management plans involving a different set of hypothetical dam removals delivering different levels of each ecosystem service. Figure 2 shows an example choice set. The watershed and subwatershed choice sets follow exactly the same experimental design, differing only in the levels.

3.4 Implementation and sample statistics

We opted for a mail survey based on reports of low broadband penetration in the study area. We sampled the names and addresses of 1,650 residents of the subwatershed from a mailing list purchased from the marketing firm infoUSA®, randomly assigning half to receive the watershed treatment and half to receive the subwatershed treatment. Thus, the sample frame is held constant between treatments. We used five points of contact via postal mail (Dillman et al., 2014), collecting data between September and November of 2018. Out of 1,310 deliverable surveys we received 448 that were at least partially completed, for a response rate of 34%. We dropped individuals who chose the Status Quo alternative in all six choice questions and indicated their reason for doing so as either disagreement with the payment vehicle or a belief that they should not be asked to pay for damage they did not cause

(Strazzera et al., 2003). Eliminating these protest responses along with unanswered questions left us with 2,286 choices made by 389 individuals. Randomization appeared to work. Responses were split approximately evenly between the two scale treatments and demographics are similar across treatments (Table 2).

4. Methodology

4.1 Empirical models

We model our data using three different utility specifications, each making a different assumption about risk preferences (Table 3). Our baseline model is a nonlinear expected utility function that assumes risk neutrality. The next model introduces risk aversion by allowing the expected utility function to exhibit concavity. And our final model incorporates a self-reported measure of risk aversion that allows us to separate risk aversion from diminishing marginal utility. We describe each model in more detail below.

Many choice experiment studies incorporating explicit risk or uncertainty have included an expected utility model, either as their main model (Faccioli et al., 2019; Makriyannis et al., 2018) or as a baseline for comparison (Glenk and Colombo, 2013; Lundhede et al., 2015; Roberts et al., 2008). We do the latter, assuming the standard random utility framework and setting v_i , the observable portion of utility from alternative i , as a linear function of the attributes. We refer to this as our baseline, or linear expected utility (LEU) model.

$$v_i = \beta_{SQ} + \beta_{salm}\pi_i x_{salm_i} + \beta_{hyd}x_{hyd_i} + \beta_{her}x_{her_i} + \beta_{lake}x_{lake_i} + \beta_{cost}x_{cost_i} \quad (3)$$

In this LEU model, we assume survey respondents make an expected utility calculation, multiplying the number of salmon that can be supported (x_{salm}) by the probability that salmon will reach that number (π). Inherent in this specification is that respondents are risk neutral. In order to easily identify the differences in subsequent models, we drop the i subscripts and collapse the remaining k attributes into a summation, so that equation (3) becomes the following.

$$v = \beta_{salm}\pi x_{salm} + \sum_k \beta_k x_k \quad (4)$$

Our second specification is a nonlinear expected utility (NLEU) model, which introduces a form of risk aversion that is confounded with diminishing marginal utility. We specify the following, which is similar to others used in choice experiments (Glenk and Colombo, 2013; Holzer and McConnell, 2017) but with a slightly simpler functional form.

$$v = \pi x_{salm}^{1-r} + \sum_k \beta_k x_k \quad (5)$$

In model (5) we estimate r , the coefficient of relative risk aversion, with $0 < r < 1$ implying both risk aversion and diminishing marginal utility. Their inseparability arises from Jensen's Inequality (i.e., $u' > 0, u'' < 0 \Rightarrow u(E[x]) > E[u(x)]$). If diminishing marginal utility is a source of scale dependency, the NLEU model may be useful to reduce it. At the same time, if we wish to analyze risk aversion itself as it relates to scale dependency, model (5) is insufficient.

One option to separate risk aversion from diminishing marginal utility is to incorporate nonlinear probability weighting, an element of prospect theory (Kahneman and Tversky, 1979).

There is some evidence that people exhibit the familiar S-shaped weighting curve in environmental risk settings, although the degree of weighting varies with context (Bartczak et al., 2015; Glenk and Colombo, 2013; Riddell, 2012; Roberts et al., 2008). Probabilistic risk aversion is attractive for behavioral reasons and the fact that it is separable from diminishing marginal utility, but the welfare implications are unclear (Holzer and McConnell, 2017). For this reason we consider an alternative.

To consider risk aversion separately from the shape of our utility function, we use a context free self-reported measure developed by Dohmen (2011) and shown to correlate with traditional experimental elicitations (Naranjo et al., 2019). We asked survey respondents the following question: *“How do you see yourself? Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?”* Then we asked them to choose a number between 0 (*“not at all willing to take risks”*) and 10 (*“very willing to take risks”*). To improve statistical power and improve interpretation, we collapse and recode into a self-reported risk measure R_{SR} , which can take a value of 0 (least risk averse), 1, or 2 (most risk averse). For our final utility specification we return to our baseline linear utility model but include R_{SR} as an interaction term. We refer to this as our self-reported risk aversion (SRRA) model.

$$v = \beta_s \pi x_s + \beta_{SR} R_{SR} \pi x_s + \sum_k \beta_k x_k \quad (6)$$

In equation (6), β_{SR} estimates the utility discount associated with increasing levels of risk aversion. This is similar to the setup used by Petrolia (2016) to model preferences for risky seafood. We summarize the three models in Table 3.

4.2 Estimation details and hypothesis tests

We split the data by scale treatment and estimate the LEU, NLEU, and SRRA specifications using mixed (random parameter) logit to account for the panel structure of the data and allow for individual taste heterogeneity. We specify all parameters as following normal distributions, except for β_{cost} , which we specify as fixed. We use the six resulting sets of parameter estimates (three models estimated using data from each of the two scale treatments) in subsequent welfare analyses. Our hypothesis tests consist of comparing parameters and welfare estimates between models and scales.

We consider the change in welfare from two restoration scenarios: fully restoring Atlantic salmon habitat in the subwatershed and fully restoring Atlantic salmon habitat in the watershed, while holding the levels of all other attributes constant. Following Hanemann (1984), we calculate an average welfare change as follows:

$$CS = -\frac{1}{\beta_{cost}} [\ln \sum_n \exp v_n^1 - \ln \sum_n \exp v_n^0] \quad (7)$$

where CS is the compensating surplus welfare measure and v_n^0 and v_n^1 are the n th residents' indirect utility before and after restoration, respectively. We construct 95% confidence intervals using the Krinsky and Robb (1986) technique.

Hypothesis 1 is a test for scale dependency, or a difference in implied prices by scale treatment. We expect the subwatershed treatment (*SUB*) to yield higher marginal WTP for Atlantic salmon restoration compared to the watershed treatment (*WS*).

$$H1: \quad mWTP^{SUB} > mWTP^{WS}$$

Hypothesis 2 tests whether risk aversion is stronger at the watershed scale than at the subwatershed scale. For this we turn to coefficients from the two models that incorporate risk aversion, NLEU and SRRA. For the NLEU model we can compare values of r , the Arrow-Pratt coefficient of relative risk aversion, with a higher value indicating a higher degree of risk aversion. For the SRRA model, we can compare estimates of β_{SR} , the marginal disutility associated with individuals who self-report higher risk aversion.

$$H2: \quad r^{WS} > r^{SUB} \text{ and } \beta_{SR}^{WS} > \beta_{SR}^{SUB}$$

Hypothesis 3 tests whether models that incorporate risk aversion can reduce scale dependency in order to reduce transfer errors. Support for Hypothesis 1 suggests the risk neutral (LEU) model results in between-scale transfer errors. Here we test whether a risk aversion model (NLEU) can address that problem. Transfer errors are eliminated if the welfare estimate for restoring salmon habitat throughout the subwatershed does not differ by scale treatment.¹

$$H3: \quad Welfare(SUB)^{SUB} = Welfare(SUB)^{WS}$$

Hypothesis 4 tests whether risk aversion becoming more important at the larger scale (Hypothesis 2) contributes to scope response. Since the subwatershed is

¹ We do not attempt to transfer welfare estimates in the other direction (i.e., from the subwatershed to the watershed), since the watershed levels are outside the range elicited in the subwatershed treatment.

embedded geographically within the watershed, and the respective quantities of Atlantic salmon are numerically embedded, the value of restoring salmon should be non-declining as we increase the restoration scale from subwatershed to watershed. We define scope response as the ratio of watershed (WS) to subwatershed (SUB) welfare estimates ($SCOPE = WS/SUB$) and characterize the results as follows: $SCOPE > 1$ indicates ‘strong’ scope; $SCOPE = 1$ indicates ‘weak’ scope; and $SCOPE < 1$ indicates scope ‘failure’. Since we wish to investigate risk aversion as a determinant of scope response separate from diminishing marginal utility, we avoid the nonlinear model (NLEU) and focus on our self-reported measure (SRRA). Using the SRRA model we calculate $SCOPE$ with the three possible values of self-reported risk aversion ($R_{SR} = 0, 1, \text{ or } 2$), where 2 indicates those with the highest degree of risk aversion. If risk aversion attenuates scope response, we would expect the following pattern.

$$H4: \quad SCOPE(R_{SR} = 0) > SCOPE(R_{SR} = 1) > SCOPE(R_{SR} = 2)$$

5. Results

We report estimated means and standard deviations for the random parameters and point estimates for the cost parameter from the three utility specifications at each of the two scale treatments (Table 4). All coefficients are in the expected direction and with one exception, all are significant with at least 90% confidence. In general, respondents prefer some new dam management plan to maintaining the status quo. They prefer plans that deliver more habitat for Atlantic salmon and river herring, while not sacrificing hydropower capacity or lakefront shoreline. The SRRA model

performs best and NLEU performs worst in terms of log likelihood, AIC, and Mcfadden's pseudo R^2 at both scales.

We note a few general points of comparison between scale treatments. First, scale treatment differences are attribute-specific. For example, marginal utility of income (i.e., the cost parameter) and the marginal utility of hydropower are virtually unaffected by scale the treatment. Point estimates of marginal utility for all other attributes are of greater magnitude at the subwatershed scale. The coefficients for fish (i.e., Atlantic salmon and herring) are about twice as large at the subwatershed scale, while the coefficient for lake shoreline is about six times as large at that scale. Since the cost parameter is consistent, implied prices follow the same pattern.

In general, there is greater preference heterogeneity at the watershed treatment, evidenced by the standard deviation estimates. Both risk aversion models exhibit statistically significant deviation around the mean only at the watershed scale, providing an initial indication that risk preferences differ by scale. The nonlinear model (NLEU) yields standard deviation estimates for nearly all parameters that are smaller in magnitude than the other two models. Finally, the standard deviation of salmon decreases when we allow marginal utility to vary by degree of self-reported risk aversion (SRRA). Next we report the results of each hypothesis test.

5.1 Scale dependency

We find support for Hypothesis 1, that implied prices for an explicitly uncertain attribute exhibit scale dependency when we assume risk neutrality. Using

the LEU model results (Table 4) we calculate marginal WTP for habitat restoration to support Atlantic salmon and construct 95% confidence intervals using Krinsky and Robb (1986) technique. The subwatershed treatment yields an average implied price of 24.59 (21.47 – 27.81) USD per person per year. With the watershed treatment, the average implied price of salmon drops to 9.55 (8.68 – 10.52) USD per person annually.

5.2 Environmental risk aversion and scale

We find support for Hypothesis 2, that people are more averse to environmental risk at the larger scale. First, comparing point estimates for the coefficients of relative risk aversion (r in our NLEU model) we found 0.923 and 0.694 at the watershed and subwatershed scales, respectively (Table 4). The higher value of r indicates a larger degree of environmental risk aversion at the watershed scale. However, since r simultaneously determines concavity of the utility function, the higher value also indicates that marginal utility diminishes more quickly at the watershed scale. Additional evidence, unconfounded with diminishing marginal utility, comes from looking at the adjustments to marginal utility associated with a higher degree of self-reported risk aversion (β_{SR} in our SRRA model). At the subwatershed scale, β_{SR} is -0.036 and not statistically significant. In contrast, β_{SR} is greater in magnitude (-0.052) and 95% statistically significant at the watershed scale. Taken together, these results suggest a direct relationship between environmental risk aversion and scale.

5.3 Benefit transfer

We find limited support for Hypothesis 3, which tests the ability of a nonlinear risk aversion model (NLEU) to reduce between-scale transfer errors. We calculate the errors associated with applying parameters derived from the watershed scale treatment to value restoring salmon to the subwatershed (Table 5). To explore the effect of heterogeneous risk preferences, we present the results using two subsamples in addition to our full sample. Of 389 total respondents in our full sample, 203 indicated disagreement with the following statement: *“If one of the particular projects is chosen, the benefits will be delivered in the amounts predicted”*. The second subsample is composed of the 237 individuals who expressed agreement with the following consequentiality statement: *“The results of this survey will influence management of the Penobscot Watershed.”* Both groups perceive some additional source of uncertainty beyond what is explicit in the survey which influence their risk preference (e.g., distrust of scientific modeling or political uncertainty). Full mixed logit results for the two subsamples are in Tables 6 and 7.

The linear model (LEU) consistently underestimates benefits when the watershed treatment is applied to value subwatershed restoration. Using the full sample, the average value of restoring salmon to the subwatershed derived from the linear model is \$179.41. Using the watershed parameters, the welfare estimate for the same restoration is \$69.76, an underestimation of 61.1%. The linear model yields similar results for the two subsamples. The watershed treatment underestimates the value of subwatershed restoration by 57.2% for the uncertainty subsample and 66.5%

for the consequentiality subsample. These results from the linear model are as expected, given the scale dependency of marginal WTP (Hypothesis 1).

The nonlinear risk aversion model (NLEU) reduces transfer errors compared to the linear model (LEU) for the two subsamples, but not the full sample. For the subsample perceiving additional uncertainty, the underestimation of benefits from the watershed treatment is reduced to 21.1%. Further, the estimates generated by the watershed and subwatershed treatments for this subsample are statistically identical, indicating that the NLEU model effectively eliminated transfer errors for this subsample. For the consequentiality subsample the reduction in error is more modest (48.8%), and the magnitude of error increases (81.6%) when the full sample is utilized.

5.4 Scope response

We find evidence supporting Hypothesis 4, that risk aversion becoming more important at the larger scale (Hypothesis 2) attenuates scope response. In the previous section we focused on welfare estimates for restoring salmon to the subwatershed. Now we introduce a second scenario: restoring salmon to the entire watershed. We perform our external (between-sample) scope test by comparing the subwatershed welfare derived from the subwatershed treatment to the watershed welfare derived from the watershed treatment. We present baseline results using the linear model (LEU) and look for the effect of risk aversion using the self-reported measure (SRRA) to avoid confounding with diminishing marginal utility (Figure 3).

Baseline analysis using our risk neutral (LEU) model suggests the presence of weak scope. Point estimates for the average value of restoring salmon to the subwatershed and watershed are \$179.41 and \$203.47, respectively. However, since the estimates are not statistically different, we conclude that welfare is not strictly increasing with the scale of restoration. Hence our baseline characterization of weak scope.

Results from the self-reported risk aversion (SRRA) model suggest that the assumption of risk neutrality in our baseline (LEU) analysis masks some heterogeneity. When $R_{SR} = 0$, the subwatershed and watershed restorations are valued at \$219.83 and \$328.48 respectively, using the appropriate model scale parameters. These estimates are statistically different, indicating that people reporting the least amount of risk aversion exhibit a strong external scope response. When $R_{SR} = 1$ we see nearly identical estimates for restoring the subwatershed and watershed, evidence of weak scope. And finally, for people reporting the highest degree of risk aversion ($R_{SR} = 2$), we see external scope failure. For these individuals, the watershed restoration is valued at \$134.15, but the greater aversion to risk at the larger scale drives the value of watershed restoration down to a statistically different \$24.73.

6. Discussion

Our baseline results demonstrating scale dependency are consistent with prior split-sample choice experiments finding lower implied prices associated with broader scales (Bueren and Bennett, 2004; Rolfe et al., 2013; Rolfe and Windle, 2008; Schaafsma and Brouwer, 2013). We show that the scale effect extends to an ecosystem

service described with explicit uncertainty. Some studies have attempted to calculate calibration factors, which would allow value estimates derived at one scale to be easily transferred to a different scale (Bueren and Bennett, 2004; Rolfe et al., 2013). Our results introduce a couple of complications to this approach. First, scale dependency was not constant across all the attributes in our choice experiment. The implied prices for additional hydropower capacity, for example, were mostly consistent between scale treatments. In contrast, implied prices for fish and lake shoreline exhibited varying degrees of scale dependency. One possibility is that respondents perceived the projections of hydropower capacity with greater certainty compared to the other attributes (Lundhede et al., 2015), despite that we only explicitly associated uncertainty with Atlantic salmon. Implementing a scale-calibration approach to benefit transfer (Rolfe et al., 2015) may require calculating different factors for certain and uncertain attributes.

Further complicating the scale issue for uncertain attributes, our results show that a risk aversion model reduced transfer errors only for subsets of respondents who perceived different levels and potential sources of uncertainty. In particular, the nonlinear utility concept of risk aversion captured by our NLEU model eliminated scale dependency for individuals who perceived additional uncertainty beyond what was explicit in the survey. This highlights the importance of controlling for prior, or perceived uncertainty (Lundhede et al., 2015). It also suggests that respondents may have different risk preferences for different sources of uncertainty. The uncertainty we describe explicitly in the survey pertains strictly to environmental sources—salmon restoration is subject to several environmental pressures that cannot be predicted with

certainty. The additional uncertainty perceived by some respondents could be political uncertainty, modeling uncertainty, or something else altogether. Even if risk preferences are consistent, they may interact with scale in unique ways depending on the source.

We find strong evidence that environmental risk aversion in general has a direct relationship with scale. A growing literature studies risk preferences specifically in an environmental context, finding consistent evidence that people are averse to environmental risk (e.g., Bartczak et al., 2015; Faccioli et al., 2019; Riddell, 2012). Beyond that, little is known aside from the fact that welfare estimates are sensitive to different modeling treatments of risk aversion (Glenk and Colombo, 2013). We showed that the risk aversion-scale relationship known to exist in the financial domain (Binswanger, 1980; Bosch-Domènech and Silvestre, 1999; Holt and Laury, 2002) applies in an environmental context. Next we discuss how our result contributes to the ongoing scope debate.

Environmental risk aversion increasing with scale is a novel consideration for interpreting scope tests. To our knowledge, only one other choice experiment study performs a split sample test for external scope (Andersson et al., 2016). They find the value of a statistical life is sensitive to the magnitude of mortality reduction presented to survey respondents, which they interpret as scope failure indicating theoretical inconsistency. Our results, however, suggest that the scope failure may result from a positive risk aversion-scale relationship, an accepted behavioral finding for financial contexts. More generally, we contribute to the large body of work exploring potential explanations for a modest or nonexistent scope response (e.g., Carson and Mitchell,

1995; Desvousges et al., 2012; Whitehead, 2016). Our work is particularly close to that of Powe and Bateman (2004), who show in a contingent valuation setting that scope can be attenuated by the perception that larger environmental programs are riskier.

7. Conclusion

We administered a choice experiment with a split sample scale treatment in order to study environmental risk aversion and potential implications for nonmarket valuation. We set our study in the context of dam removals to restore endangered Atlantic salmon habitat at the expense of hydropower capacity. We vary the likelihood that salmon will reach the number expected by scientists, reflecting uncertainty surrounding Atlantic salmon recovery efforts. We create two versions of the survey, one framed at the Penobscot watershed scale and the other framed at an embedded subwatershed scale. Holding the sample frame constant, we randomly assign subjects to receive one of the two scale treatments. We report mixed logit results derived from each scale treatment using three different utility specifications. Our baseline analysis utilizes a linear expected utility function, which inherently assumes risk neutrality. The remaining two models make different assumptions about the origin of risk aversion. The first is a nonlinear expected utility model that assumes risk aversion arises solely from concave utility, while the second treats risk aversion as an inherent human trait that respondents self-report in the survey.

Our results show that people are more averse to larger-scale environmental risks, which attenuates scope response. Both models that incorporate risk aversion indicate people discount risk more heavily at the larger scale, a behavioral finding that has already been established for financial risks (Holt and Laury, 2002). Focusing on the implications for scope, we find that the baseline risk neutral model masks a heterogeneous scope response. Our self-reported risk aversion model results show that people reporting the lowest degree of risk aversion exhibit a strong scope response while those reporting the highest degree of risk aversion exhibit scope failure. In the only comparable split-sample test for external scope using a choice experiment, Andersson et al. (2016) report scope failure. Our results suggest that they may be reporting a false negative by failing to incorporate risk preferences. In general, we suggest the risk aversion-scale relationship as a novel factor for understanding scope response.

Additionally, we find that the implied prices for salmon, but not hydropower, depend on the scale at which they are elicited. Marginal values, therefore, cannot simply be calibrated to different scales using a common adjustment factor. This poses a problem for benefit transfer approaches that rely on harvesting implied prices from choice experiments (Rolfe et al., 2015). We found that a risk aversion model could successfully eliminate scale dependency, negating the need for scale calibration, for a subset of our respondents who perceived additional sources of uncertainty. We think an important area for future research is to isolate risk preferences for the various sources and types of environmental uncertainty.

We conclude with a few general implications for stated preference research. Namely, it seems prudent to include a simple self-assessed risk preference elicitation alongside other debriefing questions. The cognitive burden to the respondent is low, and it may be helpful to demonstrate scope responsiveness. In addition, collecting even coarse risk preference data could prove beneficial for future benefit transfer research given the relationship between risk preferences and scale. While our research shows a limited ability for risk aversion models to reduce transfer errors, risk preference data could potentially be a useful input into meta-analysis benefit transfer functions. More research is required to validate the utility of this method, however, the research cannot be performed until enough primary studies have collected risk preference data.

References

- American Rivers, 2018. American Rivers Dam Removal Database. figshare. Dataset.
- Andersson, H., Hole, A.R., Svensson, M., 2016. Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety. *J. Environ. Econ. Manag.* 75, 41–53. <https://doi.org/10.1016/j.jeem.2015.11.001>
- Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M., Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. *Fish. Res., The Scientific Basis for Management of Salmonid stocks in the British Isles* 62, 143–170. [https://doi.org/10.1016/S0165-7836\(02\)00160-1](https://doi.org/10.1016/S0165-7836(02)00160-1)
- Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA Panel on Contingent Valuation.
- Auerbach, D., Deisenroth, D., McShane, R.R., McCluney, K.E., LeRoy Poff, N., 2014. Beyond the concrete: Accounting for ecosystem services from free-flowing rivers. *Ecosyst. Serv.* 10, 1–5. <https://doi.org/10.1016/j.ecoser.2014.07.005>
- Bartczak, A., Chilton, S., Meyerhoff, J., 2015. Wildfires in Poland: The impact of risk preferences and loss aversion on environmental choices. *Ecol. Econ.* 116, 300–309. <https://doi.org/10.1016/j.ecolecon.2015.05.006>
- Bartczak, A., Mariel, P., Chilton, S., Meyerhoff, J., 2016. The impact of latent risk preferences on valuing the preservation of threatened lynx populations in Poland. *Aust. J. Agric. Resour. Econ.* 60, 284–306. <https://doi.org/10.1111/1467-8489.12123>
- Bednarek, A.T., 2001. Undamming Rivers: A Review of the Ecological Impacts of Dam Removal. *Environ. Manage.* 27, 803–814. <https://doi.org/10.1007/s002670010189>
- Binswanger, H.P., 1980. Attitudes toward Risk: Experimental Measurement in Rural India. *Am. J. Agric. Econ.* 62, 395–407. <https://doi.org/10.2307/1240194>
- Bosch-Domènech, A., Silvestre, J., 1999. Does risk aversion or attraction depend on income? An experiment. *Econ. Lett.* 65, 265–273. [https://doi.org/10.1016/S0165-1765\(99\)00154-8](https://doi.org/10.1016/S0165-1765(99)00154-8)
- Boyd, J., Krupnick, A.J., 2013. Using Ecological Production Theory to Define and Select Environmental Commodities for Nonmarket Valuation [WWW Document]. *Agric. Resour. Econ. Rev.* <https://doi.org/10.22004/ag.econ.148395>

- Boyle, K.J., Desvousges, W.H., Johnson, F.R., Dunford, R.W., Hudson, S.P., 1994. An Investigation of Part-Whole Biases in Contingent-Valuation Studies. *J. Environ. Econ. Manag.* 27, 64–83.
- Bueren, M.V., Bennett, J., 2004. Towards the development of a transferable set of value estimates for environmental attributes. *Aust. J. Agric. Resour. Econ.* 48, 1–32. <https://doi.org/10.1111/j.1467-8489.2004.t01-1-00228.x>
- Carson, R.T., Mitchell, R.C., 1995. Sequencing and Nesting in Contingent Valuation Surveys. *J. Environ. Econ. Manag.* 28, 155–173. <https://doi.org/10.1006/jeem.1995.1011>
- Desvousges, W., Mathews, K., Train, K., 2012. Adequate responsiveness to scope in contingent valuation. *Ecol. Econ., The Economics of Degrowth* 84, 121–128. <https://doi.org/10.1016/j.ecolecon.2012.10.003>
- Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th ed. John Wiley, Hoboken, NJ.
- Dohmen, T., Huffman, D., Schupp, J., Falk, A., Sunde, U., Wagner, G.G., 2011. INDIVIDUAL RISK ATTITUDES: MEASUREMENT, DETERMINANTS, AND BEHAVIORAL CONSEQUENCES. *J. Eur. Econ. Assoc.* 9, 522–550.
- Faccioli, M., Kuhfuss, L., Czajkowski, M., 2019. Stated Preferences for Conservation Policies Under Uncertainty: Insights on the Effect of Individuals' Risk Attitudes in the Environmental Domain. *Environ. Resour. Econ.* 73, 627–659. <https://doi.org/10.1007/s10640-018-0276-2>
- FERC, 2019. Complete list of active and exempt licenses.
- Fischhoff, B., Quadrel, M.J., Kamlet, M., Loewenstein, G., Dawes, R., Fischbeck, P., Klepper, S., Leland, J., Stroh, P., 1993. Embedding Effects: Stimulus Representation and Response Mode. *J. Risk Uncertain.* 6, 211–234.
- Glenk, K., Colombo, S., 2013. Modelling outcome-related risk in choice experiments. *Aust. J. Agric. Resour. Econ.* 57, 559–578. <https://doi.org/10.1111/1467-8489.12012>
- Grijalva, T., Berrens, R.P., Shaw, W.D., 2011. Species preservation versus development: An experimental investigation under uncertainty. *Ecol. Econ.* 70, 995–1005. <https://doi.org/10.1016/j.ecolecon.2011.01.002>
- Hanemann, W.M., 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *Am. J. Agric. Econ.* 66, 332–341. <https://doi.org/10.2307/1240800>

- Holt, C.A., Laury, S.K., 2002. Risk Aversion and Incentive Effects. *Am. Econ. Rev.* 92, 12.
- Holzer, J., McConnell, K., 2017. Risk Preferences and Compliance in Recreational Fisheries. *J. Assoc. Environ. Resour. Econ.* 4, S1–S35. <https://doi.org/10.1086/691976>
- Houston, R., Chadbourne, K., Lary, S., Charry, B., 2007. Geographic distribution of diadromous fish in Maine. Falmouth, Maine: U.S. Fish and Wildlife Service, Gulf of Maine Coastal Program.
- Johnston, R.J., Boyle, K.J., Adamowicz, W. (Vic), Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., Vossler, C.A., 2017. Contemporary Guidance for Stated Preference Studies. *J. Assoc. Environ. Resour. Econ.* 4, 319–405. <https://doi.org/10.1086/691697>
- Johnston, R.J., Rolfe, J., Rosenberger, Randall S., Brouwer, Roy, 2015. Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners, *The Economics of Non-Market Goods and Resources*. Springer.
- Johnston, R.J., Rosenberger, R.S., 2010. Methods, Trends and Controversies in Contemporary Benefit Transfer. *J. Econ. Surv.* 24, 479–510. <https://doi.org/10.1111/j.1467-6419.2009.00592.x>
- Kahneman, D., Knetsch, J.L., 1992. Valuing public goods: The purchase of moral satisfaction. *J. Environ. Econ. Manag.* 22, 57–70. [https://doi.org/10.1016/0095-0696\(92\)90019-S](https://doi.org/10.1016/0095-0696(92)90019-S)
- Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263–291. <https://doi.org/10.2307/1914185>
- Krinsky, I., Robb, A.L., 1986. On Approximating the Statistical Properties of Elasticities. *Rev. Econ. Stat.* 68, 715–719. <https://doi.org/10.2307/1924536>
- Limburg, K.E., Waldman, J.R., 2009. Dramatic Declines in North Atlantic Diadromous Fishes. *BioScience* 59, 955–965. <https://doi.org/10.1525/bio.2009.59.11.7>
- Lopes, A., Kipperberg, G., n.d. Diagnosing insensitivity to scope in contingent valuation.
- Lundhede, T., Jacobsen, J.B., Hanley, N., Strange, N., Thorsen, B.J., 2015. Incorporating Outcome Uncertainty and Prior Outcome Beliefs in Stated Preferences. *Land Econ.* 91, 296–316. <https://doi.org/10.3368/le.91.2.296>

- Makriyannis, C., Johnston, R.J., Whelchel, A.W., 2018. Are Choice Experiment Treatments of Outcome Uncertainty Sufficient? An Application to Climate Risk Reductions. *Agric. Resour. Econ. Rev.* 1–33.
<https://doi.org/10.1017/age.2017.27>
- Martin, E., Apse, C., 2011. Northeast Aquatic Connectivity: An Assessment of Dams on Northeastern Rivers. The Nature Conservancy, Eastern Freshwater Program.
- Mills, K.E., Pershing, A.J., Sheehan, T.F., Mountain, D., 2013. Climate and ecosystem linkages explain widespread declines in North American Atlantic salmon populations. *Glob. Change Biol.* 19, 3046–3061.
<https://doi.org/10.1111/gcb.12298>
- Naranjo, M., Alpizar, F., Martinsson, P., 2019. Alternatives for risk elicitation in the field: Evidence from coffee farmers in Costa Rica. *Environ. Dev. Discuss. Pap. Ser.*
- NOAA, 2009. Biological valuation of Atlantic salmon habitat within the Gulf of Maine Distinct Population Segment Appendix A.
- Olivero, A., Anderson, M., 2008. Northeast Aquatic Habitat Classification. Boston, MA.
- Petrolia, D.R., 2016. Risk preferences, risk perceptions, and risky food. *Food Policy* 64, 37–48. <https://doi.org/10.1016/j.foodpol.2016.09.006>
- Poff, N.L., Hart, D.D., 2002. How Dams Vary and Why It Matters for the Emerging Science of Dam Removal. *BioScience* 52, 659–668.
[https://doi.org/10.1641/0006-3568\(2002\)052\[0659:HDVAWI\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0659:HDVAWI]2.0.CO;2)
- Powe, N.A., Bateman, I.J., 2004. Investigating Insensitivity to Scope: A Split-Sample Test of Perceived Scheme Realism. *Land Econ.* 80, 258–271.
<https://doi.org/10.2307/3654742>
- Riddel, M., 2012. Comparing risk preferences over financial and environmental lotteries. *J. Risk Uncertain. N. Y.* 45, 135–157.
<http://dx.doi.org.uri.idm.oclc.org/10.1007/s11166-012-9149-1>
- Roberts, D.C., Boyer, T.A., Lusk, J.L., 2008. Preferences for environmental quality under uncertainty. *Ecol. Econ.* 66, 584–593.
<https://doi.org/10.1016/j.ecolecon.2008.05.010>
- Rolfe, J., Windle, J., 2015. Do Respondents Adjust Their Expected Utility in the Presence of an Outcome Certainty Attribute in a Choice Experiment? *Environ. Resour. Econ.* 60, 125–142. <https://doi.org/10.1007/s10640-014-9760-5>

- Rolfe, J., Windle, J., 2008. Testing for differences in benefit transfer values between state and regional frameworks. *Aust. J. Agric. Resour. Econ.* 52, 149–168. <https://doi.org/10.1111/j.1467-8489.2008.00405.x>
- Rolfe, J., Windle, J., Bennett, J., 2015. Benefit Transfer: Insights from Choice Experiments, in: *Benefit Transfer of Environmental and Resource Values: A Guide for Researchers and Practitioners*. Springer, pp. 191–208.
- Rolfe, J., Windle, J., Bennett, J., Mazur, K., 2013. Calibration of values in benefit transfer to account for variations in geographic scale and scope: comparing two choice modelling experiments.
- Rollins, K., Lyke, A., 1998. The Case for Diminishing Marginal Existence Values. *J. Environ. Econ. Manag.* 36, 324–344. <https://doi.org/10.1006/jeem.1998.1045>
- Rose, J.M., Bliemer, M.C.J., 2009. Constructing Efficient Stated Choice Experimental Designs. *Transp. Rev.* 29, 587–617. <https://doi.org/10.1080/01441640902827623>
- Roy, S.G., Uchida, E., Souza, S.P. de, Blachly, B., Fox, E., Gardner, K., Gold, A.J., Jansujwicz, J., Klein, S., McGreavy, B., Mo, W., Smith, S.M.C., Vogler, E., Wilson, K., Zydlewski, J., Hart, D., 2018. A multiscale approach to balance trade-offs among dam infrastructure, river restoration, and cost. *Proc. Natl. Acad. Sci.* 115, 12069–12074. <https://doi.org/10.1073/pnas.1807437115>
- Schaafsma, M., Brouwer, R., 2013. Testing geographical framing and substitution effects in spatial choice experiments. *J. Choice Model.* 8, 32–48. <https://doi.org/10.1016/j.jocm.2013.04.007>
- Strazzera, E., Genius, M., Scarpa, R., Hutchinson, G., 2003. The Effect of Protest Votes on the Estimates of WTP for Use Values of Recreational Sites. *Environ. Resour. Econ.* 25, 461–476. <https://doi.org/10.1023/A:1025098431440>
- United States EIA, 2019. Maine - State Energy Profile Overview - U.S. Energy Information Administration (EIA) [WWW Document]. URL <https://www.eia.gov/state/?sid=ME> (accessed 3.26.20).
- USGS, 2017. National Hydrography Dataset (NHD) Plus, Version 2.
- Whitehead, J.C., 2016. Plausible responsiveness to scope in contingent valuation. *Ecol. Econ.* 128, 17–22. <https://doi.org/10.1016/j.ecolecon.2016.03.011>

Figure 1. Geographic scale treatment

		Status Quo No Action	New Plan A	New Plan B
ATLANTIC SALMON number that can be supported by available habitat	700 salmon	5,000 salmon	2,500 salmon
LIKELIHOOD that salmon will utilize newly available habitat	- 100% chance of no change to salmon	75% 25% chance of no change to salmon	50% 50% chance of no change to salmon
HYDROPOWER number of homes powered	50 thousand homes	30 thousand homes	50 thousand homes
RIVER HERRING number that can be supported by available habitat	1 million herring	3 million herring	6 million herring
LAKE SHORELINE miles lost	0 miles lost 500 remain	30 miles lost 470 remain	15 miles lost 485 remain
ANNUAL COST TO YOU property tax surcharge, each year for 10 years	\$0	\$400	\$50
Which would you choose? <u>check one box ONLY</u>		No Action ☐	Plan A ☐	Plan B ☐

Figure 2. Sample choice set from the subwatershed treatment

Figure 3. External scope test results by level of self-reported risk aversion

Notes: Welfare estimates are for restoring Atlantic salmon habitat to the subwatershed and watershed. Levels of other attributes are held constant. Error bars represent 95% confidence intervals constructed using Krinsky and Robb procedure with 5,000 draws.

Table 1. Attribute levels for both scale treatments

Attribute Name	Levels							
	Subwatershed				Watershed			
ATLANTIC SALMON number that can be supported	700	2500	5000	8000	1500	7000	14000	22000
LIKELIHOOD that salmon utilize new habitat			25%	50%	75%	90%		
HYDROPOWER thousands of homes powered	10	30	50	70	45	90	155	185
RIVER HERRING millions that can be supported	1	3	6	9	2	5	10	16
LAKE SHORELINE miles lost	0	15	30	50	0	150	300	500
ANNUAL COST property tax surcharge	\$0	\$50	\$95	\$150	\$250	\$400		

Note: Status quo levels are in bold.

Table 2. Demographic distributions by scale treatment

		Subwatershed Treatment	Watershed Treatment
Gender	Male	49.2	49.5
	Female	50.3	50.5
Age	18-34	14.7	15.7
	35-49	13.2	19.1
	50-64	34.5	35.8
	65+	37.6	29.4
Household Income (thousands)	< 50	41.8	38.2
	50 - 99	32.9	35.5
	100 - 149	16.5	17.2
	150-199	7.1	4.8
	200+	1.7	4.3
Number of respondents		191	198

Table 3. Utility specification

Name	Risk Preference Assumption	Model
Linear Expected Utility (LEU)	Risk neutrality	$v = \beta_{salm} \pi x_{salm} + \sum_k \beta_k x_k$
Nonlinear Expected Utility (NLEU)	Risk aversion confounded with diminishing marginal utility	$v = \pi x_{salm}^{1-r} + \sum_k \beta_k x_k$
Self-reported Risk Aversion (SRRA)	Risk aversion as an inherent context-free trait	$v = \beta_s \pi x_s + \beta_{SR} R_{SR} \pi x_s + \sum_k \beta_k x_k$

Table 4. Mixed logit results

	Subwatershed			Watershed		
	LEU	NLEU	SRRA	LEU	NLEU	SRRA
	Mean					
β_{Salmon}	0.154*** (0.042)		0.186*** (0.058)	0.069*** (0.020)		0.112*** (0.029)
r		0.694*** (0.081)			0.923*** (0.108)	
$\beta_{RiskAverse}$			-0.036 (0.044)			-0.052** (0.025)
$\beta_{StatQuo}$	-2.327*** (0.400)	-0.769*** (0.253)	-1.900*** (0.409)	-1.683*** (0.322)	-1.315*** (0.231)	-1.858*** (0.326)
β_{Hydro}	0.013*** (0.003)	0.010*** (0.002)	0.013*** (0.003)	0.011*** (0.002)	0.006*** (0.001)	0.011*** (0.002)
$\beta_{Herring}$	0.071*** (0.021)	0.062*** (0.020)	0.073*** (0.021)	0.028* (0.016)	0.027** (0.010)	0.028* (0.016)
β_{Lakes}	-1.823*** (0.472)	-1.593*** (0.368)	-1.867*** (0.455)	-0.314*** (0.062)	-0.190*** (0.032)	-0.316*** (0.067)
	Standard deviation					
β_{Salmon}	0.154* (0.088)		0.039 (0.196)	0.143*** (0.037)		0.098** (0.045)
r		0.114 (0.112)			0.409*** (0.087)	
$\beta_{RiskAverse}$			0.069 (0.085)			0.075** (0.036)
$\beta_{StatQuo}$	3.680*** (0.361)	1.867*** (0.256)	4.683*** (0.573)	2.911*** (0.361)	1.607*** (0.173)	2.915*** (0.364)
β_{Hydro}	0.011** (0.005)	0.000 (0.007)	0.013*** (0.005)	0.014*** (0.002)	0.008*** (0.001)	0.014*** (0.002)
$\beta_{Herring}$	0.029 (0.034)	0.001 (0.033)	0.013 (0.032)	0.105*** (0.024)	0.038*** (0.014)	0.101*** (0.024)
β_{Lakes}	2.535*** (0.684)	0.946** (0.403)	2.316*** (0.656)	0.420*** (0.079)	0.143*** (0.036)	0.442*** (0.076)
	Fixed parameters					
β_{Cost}	-0.006*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.007*** (0.001)	-0.006*** (0.000)	-0.007*** (0.001)
N (obs./ind.)	1112/191	1112/191	1112/191	1174/198	1174/198	1174/198
LL	-763.55	-912.90	-749.69	-886.85	-957.76	-879.85
AIC	1549.1	1847.8	1525.4	1795.7	1937.5	1785.7
Pseudo R-sq	0.3750	0.2527	0.3863	0.3124	0.2574	0.3178

Note: *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Table 5. Transfer errors from applying watershed treatment results to value restoration of Atlantic salmon habitat throughout the subwatershed

SAMPLE	MODEL	TREATMENT		% ERROR WS→SUB
		SUB	WS	
Full sample (N = 389)	Linear Expected Utility (LEU)	179.41 (156.60; 203.59)	69.76 (63.07; 76.90)	-61.1%
	Nonlinear Expected Utility	180.30 (121.70; 247.78)	33.24 (-13.30; 89.05)	-81.6%
Respondents who perceive additional uncertainty (N = 203)	Linear Expected Utility (LEU)	92.37 (62.48; 92.37)	39.50 (28.52; 51.11)	-57.2%
	Nonlinear Expected Utility	132.70 (89.96; 183.79)	104.74 (69.29; 143.91)	-21.1%
Respondents who perceive survey consequentiality (N = 237)	Linear Expected Utility (LEU)	226.09 (194.95; 258.46)	75.82 (67.68; 85.53)	-66.5%
	Nonlinear Expected Utility	122.77 (88.25; 161.57)	62.89 (40.73; 87.01)	-48.8%

Notes: Welfare estimates are in USD per person per year. Levels of other attributes are held constant. 95% confidence intervals are in parentheses calculated using Krinsky and Robb procedure with 5,000 draws.

Table 6. Mixed logit results for subsample who perceived additional uncertainty

	Subwatershed		Watershed	
	LEU	NLEU	LEU	NLEU
Mean				
β_{Salmon}	0.102*		0.039*	
	(0.056)		(0.021)	
r		0.776***		0.773***
		(0.120)		(0.070)
$\beta_{StatQuo}$	-1.982***	-2.019***	-1.303***	-1.007**
	(0.681)	(0.587)	(0.447)	(0.400)
β_{Hydro}	0.017***	0.012***	0.006***	0.005***
	(0.005)	(0.004)	(0.002)	(0.001)
$\beta_{Herring}$	0.057*	0.075**	0.013	0.015
	(0.033)	(0.038)	(0.018)	(0.015)
β_{Lakes}	-1.833**	-1.729***	-0.238***	-0.167***
	(0.719)	(0.522)	(0.072)	(0.055)
Standard deviation				
β_{Salmon}	0.025		0.059	
	(0.302)		(0.058)	
r		0.266*		0.037
		(0.150)		(0.110)
$\beta_{StatQuo}$	5.238***	3.031***	3.260***	2.753***
	(0.880)	(0.579)	(0.493)	(0.436)
β_{Hydro}	0.015**	0.008*	0.010***	0.001
	(0.007)	(0.005)	(0.003)	(0.003)
$\beta_{Herring}$	0.032	0.026	0.051	0.015
	(0.057)	(0.150)	(0.040)	(0.020)
β_{Lakes}	2.506***	1.199*	0.353***	0.288***
	(0.952)	(0.623)	(0.090)	(0.052)
Fixed parameters				
β_{Cost}	-0.008***	-0.006***	-0.007***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
N (obs./ind.)	598/103	598/103	592/100	592/100
LL	-360.18	-474.43	-433.27	-464.24
AIC	742.4	970.9	888.5	950.5
Pseudo R-sq	0.4517	0.3336	0.3338	0.2862

Note: *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Table 7. Mixed logit results for subsample who perceived the survey as consequential

	Subwatershed		Watershed	
	LEU	NLEU	LEU	NLEU
Mean				
β_{Salmon}	0.160*** (0.049)		0.089*** (0.034)	
r		0.776*** (0.120)		0.852*** (0.090)
$\beta_{StatQuo}$	-2.506*** (0.613)	-2.019*** (0.587)	-2.638*** (0.466)	-2.145*** (0.315)
β_{Hydro}	0.010*** (0.003)	0.012*** (0.004)	0.017*** (0.003)	0.008*** (0.002)
$\beta_{Herring}$	0.085*** (0.026)	0.075** (0.038)	0.032 (0.025)	0.012 (0.017)
β_{Lakes}	-1.764*** (0.515)	-1.729*** (0.522)	-0.447*** (0.106)	-0.304*** (0.055)
Standard deviation				
β_{Salmon}	0.093 (0.166)		0.215*** (0.056)	
r		0.266* (0.150)		0.565*** (0.083)
$\beta_{StatQuo}$	3.400*** (0.361)	3.031*** (0.579)	2.828*** (0.469)	1.584*** (0.253)
β_{Hydro}	0.006 (0.008)	0.008* (0.005)	0.020*** (0.004)	0.012*** (0.002)
$\beta_{Herring}$	0.004 (0.034)	0.026 (0.150)	0.151*** (0.035)	0.105*** (0.025)
β_{Lakes}	1.664* (0.905)	1.199* (0.623)	0.680*** (0.133)	0.246*** (0.064)
Fixed parameters				
β_{Cost}	-0.005*** (0.001)	-0.006*** (0.001)	-0.009*** (0.001)	-0.007*** (0.001)
N (obs./ind.)	648/110	648/110	750/127	750/127
LL	-451.51	-474.43	-551.31	-589.49
AIC	925.0	970.9	1124.6	1201.0
Pseudo R-sq	0.3658	0.3336	0.3309	0.2846

Note: *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Environmental federalism in river restoration

(To be submitted to Land Economics)

Ben Blachly¹ and Emi Uchida

¹ Ben Blachly is corresponding author and is a PhD candidate in the Department of Environmental and Natural Resource Economics, University of Rhode Island (ben_blachly@uri.edu). Emi Uchida is a Professor in the Department of Environmental and Natural Resource Economics, University of Rhode Island. Support for this project is provided by the National Science Foundation's Research Infrastructure Improvement NSF #IAA-1539071. We thank Erik Martin from *The Nature Conservancy* for his crucial assistance building our dataset.

ABSTRACT

Environmental federalism studies the tradeoffs between centralized and decentralized environmental policies. Ecological restoration is a large and growing sector of the economy, but it is unstudied from an environmental federalism perspective. The implications are important, because coordinating restoration projects across larger geographies increases the efficiency of ecosystem service production. We study the effects of decentralized decisions about dam removals, an increasingly popular approach to river restoration. Specifically, we ask whether state border proximity impacts the likelihood that a dam is chosen for removal. We constructed a spatially explicit dataset of removed and existing dams in 13 northeastern states. We use binary and continuous matching to improve covariate overlap, and we employ an instrumental variable to address endogeneity. Our results suggest dams located immediately upstream of a state border are about 7 percentage points more likely to be removed than dams with no border downstream. The effect becomes stronger for dams located farther upstream of a border. Each standard deviation of additional upstream distance from a border increases removal likelihood by between 5.8 and 7.0 percentage points. The qualitative consistency of our estimates across models provides evidence that states respond to individual incentives, suggesting efficiency gains from larger scale coordination may go unrealized without formal cooperation.

1. Introduction

Environmental federalism applies theories from fiscal federalism to environmental policy. The central result is that local policies can better match the tastes of local constituencies, but centralized policies increase efficiency by internalizing any externalities (Besley and Coate, 2003; Oates, 1972). For example, spillovers can cause underprovision or spatial misallocation of public goods when they are provided by local governments (Bloch and Zenginobuz, 2007; Braid, 2010). Such could be the case with certain ecosystem services, including most regulatory and cultural services, which are public in nature (Costanza, 2008). Despite a large and growing ecological restoration sector of the US economy (BenDor et al., 2015), the environmental federalism consequences for ecosystem service provision are unstudied.

This study takes a first look at environmental federalism in an ecological restoration context, examining whether states consider downstream border proximity when selecting river reaches to restore by dam removal. Dam removal is gaining popularity as an effective approach to restoring river ecosystem services (Bednarek, 2001). The productive efficiency benefits from centralizing dam removal decisions through large scale (i.e., regional or basin) optimization are well-documented (Neeson et al., 2015; Roy et al., 2018). In practice, however, dam removal decisions are generally not coordinated across sites (Magilligan et al., 2016). We empirically test whether state-level restoration outcomes are misaligned with broader optimization efforts.

Our application contributes to the empirical environmental federalism literature, which has focused mainly on air and water quality (i.e., pollution and abatement). In general, previous literature has confirmed the theoretical result that in the presence of spillovers, decentralized policies lead to suboptimal outcomes. Decentralization has been found to be associated with degraded water quality near international (Sigman, 2002) and subnational (Lipscomb and Mobarak, 2017; Sigman, 2005) borders. Air and water pollution abatement are lower when out of state constituencies receive the benefits (Gray and Shadbegian, 2004). Furthermore, countries overexploit their water resources when the costs are borne by a downstream neighbor (Olmstead and Sigman, 2015). Our research contributes to this strand of literature in a novel setting—environmental restoration.

To test environmental federalism in dam removal decisions, we constructed a spatially explicit dataset of all existing and removed dams in 13 states in the northeastern US. Taking a quasi-experimental approach, we categorize dams upstream of a state border as treated, and those with no state border downstream as controls. Noting severe misbalance between the treated and control units, we preprocess the data using covariate matching with treatment defined as binary (Ho et al., 2007) and continuous (Fong et al., 2018). We create two matched datasets, one defining treatment as distance to downstream border (continuous) and the other using presence of a downstream border (binary). We model the determinants of dam removal on our raw, binary, and continuously matched datasets, employing an instrumental variable to address endogeneity from nonrandom siting of dams relative to state borders.

We find consistent evidence that states respond to individual incentives when choosing dam removal projects. Dam removals are more likely upstream of state borders, and they become even more likely farther upstream. Our results indicate that a dam located immediately upstream of a border is about 7 percentage points more likely to be removed than a similar dam with no border downstream. Additionally, dam removal likelihood increases by 0.5 percentage points for every 10 km of downstream distance between the dam and a border. The consistency of our estimates across matching methods suggests that choosing between continuous and dichotomized matching is relatively unimportant for the purpose of data preprocessing.

2. Background

The true number of dams in the US is unknown, but estimates range from about 75,000 structures large enough to meet National Inventory of Dams criteria (Graf, 1999) to over two million when smaller barriers are included (Smith et al., 2002). Dams are physical barriers that alter the magnitude and timing of river flows to serve a purpose: provide hydropower, flood control, water supply, navigation, or recreation. But the needs of society are evolving, and the ecosystem service benefits from free-flowing rivers are increasingly clear (Auerbach et al., 2014). Dam removal has proven to be an effective tool for restoring some of these benefits, especially recovering native fish species (Bednarek, 2001).

To date there have been nearly 4,000 purposeful dam removals globally (Ding et al., 2019), and over 1,400 in the US alone (American Rivers, 2018). Grabowski et al. (2018) analyzed trends in dam removal and predicted that by 2050 the US will remove between 4,000 and 36,000 dams. Since different dam removals will deliver different levels of different ecosystem services (Poff and Hart, 2002), an important question is which removal projects should receive priority (Doyle et al., 2003).

Dam removal budgets can be utilized most efficiently by coordinating across larger spatial scales (Neeson et al., 2015; Roy et al., 2018), and researchers have developed various models to optimize dam removal decisions. Some consider dams simultaneously with culverts (Fitzpatrick and Neeson, 2018; Martin and Levine, 2017), while others optimize based on multiple ecosystem services (Roy et al., 2018). Others attempt to incorporate hydrologic modelling to improve the accuracy of predicted changes (Ishiyama et al., 2018) or economic costs to consider limited budgets (Kraft et al., 2019; Neeson et al., 2015).

Despite the availability of increasingly sophisticated models, barriers remain to a large- scale coordinated approach (Magilligan et al., 2016). For instance, well-planned and high-value dam removal projects can fail on the basis of local opposition (Fox et al., 2016). More recently, researchers have begun searching for solutions to the issue of public opposition using participatory exercises (Diessner et al., 2020), bargaining techniques (Weir et al., 2020), and multicriteria decision-making analysis (Fox et al., 2019). However, implementation of these techniques in an actual decision-making context is still at its infancy. In this paper we take a step back to examine a more fundamental barrier to coordination: strategic behavior by states.

States have the ability to steer both internal and external funds toward favored projects. American Rivers, a national non-profit organization and leading proponent of dam removal, published a guide outlining the steps involved in a typical removal project (Graber et al., 2015). Between feasibility studies, permitting, engineering, and deconstruction, projects typically require between \$50,000 and \$500,000. Funding is usually received from multiple sources. And while many major funding sources are federal in scope (e.g., the National Oceanic and Atmospheric Association), they favor projects with demonstrated state-level support (Graber et al., 2015). States can influence removals, but it is only inefficient if states prioritize projects differently than stakeholders with a broader scope. In the next section we motivate why this might be the case.

3. Environmental federalism and dam removal decisions

State-level incentives lead to dam removal decisions that may differ from regional or basin-scale optimization efforts. The realized benefits of dam removal are site-dependent, but functional network size (the number of barrier-free river miles up and downstream of a dam) is a commonly used metric to compare potential removals. When one dam is removed, the functional networks of adjacent dams increase. Those dams become more attractive removal candidates, increasing the likelihood that they will receive competitive external funds in the future. If we consider multiple time periods, the spillover benefits from one dam removal can extend well beyond immediately adjacent dams. If states wish to capture more external funds in the future

for river restoration, they may favor projects located farther upstream of state borders. This setting leads to the central hypothesis of this study:

The probability of dam removal increases with downstream distance to a state border.

After carefully controlling for other important determinants of dam removal, empirical evidence in favor of this hypothesis would be strong evidence for state-level strategizing. Nevertheless, we firmly ground our hypothesis in the literature.

Fiscal federalism, the branch of public economics studying the appropriate level of government to provide a service, provides a theoretical basis. In the case of public good provision, a key finding is that the optimal extent of decentralization depends on the extent of spillovers and constituent taste heterogeneity (Besley and Coate, 2003; Oates, 1972). By assuming away taste heterogeneity, studies have arrived at interesting results from a game theory perspective. In the presence of spillovers, equilibrium provision locations and levels differ from the social optimum (Bloch and Zenginobuz, 2007; Braid, 2010). In an environmental context, empirical evidence of fiscal federalism inefficiencies are collected in a strand of literature known as environmental federalism.

Our study examines the consequences of environmental federalism for hydrological ecosystem restoration. To date, the literature has focused mainly on air and water pollution. The strongest results are demonstrated along rivers, where directional flow is straightforward to model. For instance, Gray and Shadbegian (2004) found that net pollution was higher downstream of pulp and paper mills located upstream of state borders. Another series of studies found similar results for general

water pollution near state (Sigman, 2005) and international (Sigman, 2002) borders. Lipscomb and Mobarak (2017) exploit changing county boundaries in Brazil to implement a quasi-experimental design, finding that water quality deteriorates with border proximity. In a slight variation on the theme, Olmstead and Sigman (2015) show that countries tend to overexploit their water resources when the costs are borne by a downstream country. Specifically, countries are more likely to site large dams upstream of international borders. In contrast, we study the effect of borders on dam *removals*. Thus, we contribute to the overall literature by providing empirical evidence of strategic behavior by states in an unexplored setting, i.e., ecosystem restoration.

We note that in addition to our border proximity hypothesis, we address a related but distinct question: whether or not dam removal is more likely upstream of a state border regardless of proximity. If removal costs or risks are concentrated downstream, such as when a dam removal releases contaminated sediment or excessive nutrients (Gold et al., 2016), removals may be more likely if the costs will eventually be passed to a downstream neighbor. However, given the site specificity of removal outcomes (Poff and Hart, 2002), the expected result is ambiguous.

4. Data

We constructed a database of existing and removed dams in 13 northeastern US states. We started with a list of existing dams compiled by *The Nature Conservancy* sourced from state databases (Martin and Apse, 2011), and a list of dams removed between 1990 and 2017 compiled by American Rivers (American Rivers,

2018). Because most dam records are kept at the state level, and each state has different standards, there is a large degree of inconsistency in the scope of data available for each dam. We required only that dams were spatially explicit (i.e., latitude and longitude was available). Following Martin and Levine (2017), we loaded individual dam points into ArcGIS and snapped them to the nearest flowline available on the Medium Resolution National Hydrography Dataset Plus, Version 2 (NHDPlus v2) (Horizon Systems, 2012). We used a snapping tolerance of 100 meters to avoid erroneous associations, and kept the remaining dams that fell in USGS hydrologic unit code (HUC) regions 01 and 02.

Our final data set contains 11,601 dams, of which 395 (3.4%) are identified as removed. Consistent with global and national data, our sample indicates an upward trend in dam removals (Figure 1). Existing and removed dams are unevenly distributed between states (Figure 2). Delaware has the fewest existing dams (91) and New York has the most (1,972). About half of the removed dams in our sample (197) are concentrated in one state, Pennsylvania. Massachusetts has the second highest number of removals (35).

For each dam, we calculated a suite of physical and ecological attributes that are likely to be important determinants of removal status using custom ArcGIS tools developed by Martin and Apse (2011). We associated each dam with census tract-level demographics (U.S. Census Bureau, 2017), acknowledging that individual dam removal efforts are sometimes derailed by local public opposition (Fox et al., 2016). We categorized potential removal determinants as physical, ecological, and

demographic controls¹ (Table 1). We identify dams upstream of a state border as ‘treated’, and those with no state border downstream as ‘controls’. Of 11,601 dams in our full sample, 6,681 (57.6%) are located upstream of a state border. Finally, we calculate our continuous treatment variable, downstream distance to a state border, using ArcGIS and R.

We compare summary statistics between existing and removed dams using t-tests (Table 1, first three columns). On average, removed dams have fewer downstream barriers than existing dams, though they are roughly equidistant from the river mouth. Removed dams tend to be on higher stream order (i.e., relatively larger) rivers that are less sloped than existing dams. Removed dams have larger functional networks, exist in watersheds with greater overall species richness, and have a higher number of anadromous species documented downstream than existing dams. Finally, removed dams are more likely to be in densely-populated and lower-income census tracts. The differences are generally consistent with the dam removal literature, which finds removed dams are unrepresentative of the existing dam population (Grabowski et al., 2018; Magilligan et al., 2016).

We also use t-tests to compare characteristics of dams that are upstream of a state border (treated dams) and dams that have no state border along their downstream flowpath (control dams) (Table 1, last three columns). Statistically significant differences exist between treatment and control groups across all but one variable, the proportion falling on conserved land. The systematic differences between treatment

¹ We do not control for the expected cost of removal, since existing removal cost models are based on dam size (height and/or length) that is not available consistently in our data. However, within our set of physical controls are geologic features that are related to dam size, and thus expected removal cost.

and control groups are likely driven by their spatial distribution, since treated dams are much more likely to be located away from the coast (Figure 3). The observed differences between treatment and control motivate our empirical approach, which we outline in the next section.

5. Empirical strategy

We use logistic regressions to model the determinants of dam removal. In particular, we are interested in clean identification of a state border proximity effect, which would indicate the presence of an environmental federalism story. We employ two strategies. First, we preprocess the data to reduce model dependency using two different matching approaches. Then we employ an instrumental variable for border proximity to address endogeneity.

5.1 Logit model

We observe y , the binary removal status of each dam.

$$y_i = \begin{cases} 1 & \text{if dam } i \text{ is removed} \\ 0 & \text{if dam } i \text{ is existing} \end{cases}$$

We relate the conditional probability that dam i is removed, p_i , to a vector of observable features of the dam denoted by $\mathbf{x}_i$ through a function $F(\cdot)$ with a vector of parameters $\boldsymbol{\beta}$.

$$p_i = \Pr[y_i = 1|\mathbf{x}] = F(\mathbf{x}_i'\boldsymbol{\beta}) \tag{1}$$

We specify $F(\cdot)$ as the cumulative distribution function of the logistic distribution to arrive at a logit model, which can be rearranged so the log-odds ratio is linear in $\mathbf{x}_i$.

We expand below for clarity with the i subscripts dropped.

$$\ln \frac{p}{1-p} = \delta \cdot D_{US} + \gamma \cdot D_{US} \cdot Dist_{US} + \mathbf{x}'_{cont} \cdot \boldsymbol{\beta}_{cont} + \alpha_s + u \quad (2)$$

In equation (2), D_{US} is a dummy variable taking the value of 1 if the dam is upstream of a state border, 0 otherwise, and $Dist_{US}$ is the downstream distance (in km) to a state border. $\mathbf{x}_{cont}$ is a vector of physical, ecological, and demographic controls regarded as important determinants of p . We include α_s , a vector of state-level fixed effects, to control for time-invariant factors that make states more or less likely to remove dams. Finally, u is an error term. δ estimates the effect of being upstream of a border on the log-odds of removal, and γ indicates how the effect differs depending on border proximity. We report marginal effects for our main variables according to the following equation.

$$\text{Marginal effect of } D_{US} = F'(\bar{\mathbf{x}}'_T \boldsymbol{\beta}) \cdot \delta \quad (3)$$

$$Dist_{US} = F'(\bar{\mathbf{x}}'_T \boldsymbol{\beta}) \cdot \gamma$$

In equation (3), $F'(\cdot)$ is the partial derivate of equation (1) with respect to the relevant variable, D_{US} or $Dist_{US}$, and $\bar{\mathbf{x}}_T$ is a vector containing the mean covariate values for the treated group. Thus, we report marginal effects evaluated at the mean values of covariates of the treatment group.

5.2 Binary and continuous matching to preprocess the data

We use two matching methods to preprocess our data prior to model estimation (Fong et al., 2018; Ho et al., 2007). Our main variable of interest is a continuous measure of downstream distance to a state border, $Dist_{US}$. In constructing that variable, we necessarily identify which dams are upstream of a state border, D_{US} . Our data, therefore, provides an opportunity to compare model results obtained after preprocessing the data using binary and continuous matching. Despite the availability of methods to accommodate continuous treatment, it is still common in applied work to dichotomize continuous treatments and perform binary matching. In the economics literature we could find just three examples of matching using continuous treatment (Berhane et al., 2014; BIRTHAL et al., 2015; Magrini et al., 2017). A few studies report loss of information from binary matching compared to continuous (e.g., Cerulli et al., 2018; Fong et al., 2018). However, no such information loss is expected in a data preprocessing application where matching precedes model estimation. Instead we can check the consistency of results.

Our goal with both forms of matching is to reduce potential bias resulting from model dependency. As noted, significant differences exist between dams that are upstream of state borders (treated observations) and dams that have no border downstream (controls) across most of the observed variables (Table 1). These systematic differences indicate a lack of covariate balance between the two groups, which forces our model to extrapolate over the range that lacks common support. Ho et al. (2007) show how extrapolation leads to coefficient estimates that are sensitive to functional form assumptions, and how matching can reduce the sensitivity.

Given the relatively large number of potential confounders in our data, we match on a propensity score using the binary treatment framework developed by Rosenbaum and Rubin (1983) and extended to multivalued and continuous treatments by Imbens (2000). We calculate a propensity score using the scoring function that optimizes covariate balance between treated and control groups (Fong et al., 2018; Imai and Ratkovic, 2014) using the R package CBPS (Fong et al., 2017). The method diminishes potential effects from misspecifying the propensity score model for both binary (Imai and Ratkovic, 2014) and continuous (Fong et al., 2018) treatments. Then we use the R package MatchIt (Ho et al., 2011) to implement 3:1 nearest-neighbor matching with replacement. Because controls can be and are reused as matches, we calculate weights that we use in subsequent regressions.

5.3 Instrumental variable analysis

Our variable of interest, $Dist_{US}$, which measures the downstream distance between a dam and a state border, is potentially endogenous: unobservable factors may play a role in determining a dam's location relative to a state border. For instance, dams tend to be more prevalent closer to state and international borders as a result of state-level freeriding in water quantity (Olmstead and Sigman, 2015) or quality (Sigman, 2005). Since similar factors could contribute to dam removal decisions, $Dist_{US}$ likely suffers from endogeneity.

We introduce a plausibly exogenous variable, the percentage of the census tract population commuting to work out of state, as an instrument for downstream

distance to a state border. The quality of our instrument rests on two assumptions. First, the percentage of the population commuting out of state must be a strong predictor of downstream distance to a state border. We test the validity of this assumption by estimating the following first stage model and assessing the F-statistic.

$$Dist_{US} = \theta \cdot Commute + \mathbf{x}'\boldsymbol{\beta} + \epsilon \quad (4)$$

In equation (4), *Commute* is the instrument, the percentage of the population commuting to work out of state, and $\mathbf{x}$ contains all the same controls as equation (2).

The second assumption is that the percentage of the population commuting to work out of state only impacts dam removal likelihood through $Dist_{US}$. Though untestable, we consider it a reasonable and logical assumption. Since our second stage model is nonlinear (equation 2) we proceed using the two-stage residual inclusion (2SRI) method (Terza et al., 2008) and obtain standard errors via bootstrapping. If the exclusion restriction holds, we have purged the endogenous variation from our estimate for γ , the effect of state border proximity on the log-odds of dam removal.

6. Results

6.1 Covariate balance in the matched datasets

Datasets resulting from the two matching procedures (i.e., using binary (Imai and Ratkovic, 2014) versus continuously (Fong et al., 2018) defined treatment) include the same treated observations but different control groups. Of 4,920 potential control observations, binary and continuous matching utilize 2,824 and 4,479,

respectively. We compare covariate means across treatment and control groups using t-tests (Table 2). Statistically significant differences persist post-matching, and they are particularly large in the continuous treatment dataset. Nonetheless the post-matching t-statistics are a marked improvement over their pre-matching counterparts (Table 1). Aside from that, these types of hypothesis tests are inappropriate to assess balance despite their common usage (Imai et al., 2008).

The main purpose of matching is to improve covariate overlap between treatments and controls, which we assess in more detail for an important covariate. The downstream distance from a dam to the river mouth is clearly correlated with the downstream distance to a state border, if one exists. It is also a major determinant of dam removal decisions, since dams located closer to the coast will have greater benefits for sea-run fish. In our raw data the average downstream distance to the coast is 276.2 km for treated dams and 81.7 km for control dams, a large statistical difference (t-stat = 82.2). Observing the distributions (Figure 4a) clarifies the issue. There are few control dams available for matching to the treated dams that are farthest away from the coast. Binary matching eliminates the statistical difference in means (t-stat = 0.78) but it persists following continuous matching (t-stat = 41.2). However, both matching approaches provide vast improvements in distributional overlap (Figure 4b and 4c).

6.2 Regression results

Results from model (2) using the raw (unmatched) data indicate that being upstream of a state border increases the probability of removal by 1.8 percentage points, and the probability increases by 0.08 percentage points for every 10 km of additional upstream distance (Table 3 Column 3). The results are qualitatively robust to the inclusion of controls, but the control variables and state fixed effects have a noticeable impact on magnitude. As expected, excluding controls leads to an underestimation of the border proximity effect owing to correlation with treatment and the negative relationship between distance to the coast and removal likelihood. Excluding state fixed effects overestimates the effect from being upstream of a state border, likely due to the high concentration of overall removals in Pennsylvania. We report pseudo R-squared values calculated using the McKelvey and Zavoina method, which most closely mirrors the interpretation of OLS R-squared (Veall and Zimmermann, 1996). We loosely take the pseudo R-squared of 0.45 (Table 3 Column 3) to mean that model (1) explains about half the variation in dam removal likelihood.

The estimates from model (2) on matched datasets are directly comparable to the estimates on our raw data (Table 3). The results are qualitatively similar across unmatched and matched datasets, with some differences in magnitude. Both flavors of matching yield estimates for the state border effect that are larger than the unmatched estimates. The effect of having a downstream state border on dam removal probability is slightly larger for binary matching (3.8 percentage points) compared to continuous matching (2.8 percentage points). The additional effect for each 10 km in downstream distance to the border is 0.06 percentage points using binary matching and 0.09 percentage points using continuous matching. In other terms, dams that are one

standard deviation farther upstream from a state border are between 0.7 and 1.0 percentage points more likely to be removed. Pseudo R-squared is slightly higher for binary (0.51) compared to continuous matching (0.46). Overall, the two types of matching do not result in large differences in the marginal effect estimates from equation (2).

We turn now to the instrumental variable results using the unmatched, binary, and continuously matched datasets (Table 4). The first stage OLS estimates indicate that the instrument is strongly correlated with distance to state boarder. A one percentage point increase in the census tract population commuting out of state is associated with dams in the tract being between 1.6 and 1.9 km closer downstream to a state border. Statistical significance of the coefficients and high F-statistics, between 450.1 and 834.7, indicate strong correlation of the instrument with the endogenous variable.

The instrumental variable results are consistent across matching methods, and we find larger effect sizes compared to our naïve analysis. Removal probability increases by between 6.2 and 7.4 percentage points when a dam is located upstream of a state border. For those dams, each 10 km of downstream distance to a state border increases the probability of removal by between 0.5 and 0.6 percentage points. In terms of standard deviations, each standard deviation of upstream distance from a state border increases the removal likelihood by between 5.8 and 7.0 percentage points. Pseudo R-squared values are similar to other analyses, with the highest value obtained from the binary matched dataset (0.52).

7. Discussion

To put our results in context, we compare the probability of removal for dams located different distances upstream of a state border. We present the calculations using the instrumental variable results from our two matched datasets (Table 4 Columns 2 and 3). First, a dam located immediately upstream of a state border is 7.1 – 7.3 percentage points more likely to be removed than a similar dam with no downstream border. Compare that to a dam located 106.9 km upstream of a border, the mean value in our data. For this dam, removal is 5.7 – 6.5 percentage points more likely than the first dam, and 13.1 – 13.6 percentage points more likely than a similar dam with no downstream border.

The differences in probability of removal upstream of borders versus when no border is downstream are statistically significant, but the magnitudes are modest. Despite the overall trend, the small proportion of removed dams (3.4% in our data) suggests that dam removal is still in its infancy. A geospatial analysis of removed dams in six New England states concluded that removals have resulted from ‘strategic opportunism’, whereby removal advocates respond opportunistically to individual dam owners who are usually driven by economic or safety concerns (Magilligan et al., 2016). We agree generally with this characterization, but our results suggest there is more ‘strategy’ on the part of some actors, particularly at the state level.

We note, however, a few caveats and alternative explanations. Despite our best efforts to eliminate bias and isolate strategic behavior by states, we faced significant data limitations. Most notably, we lack data on the age of individual dams and their

purpose for removal. It could be that states favor dam removals upstream of borders to achieve some other, nonstrategic goal. We do not know, for example, whether dam building was spatially and temporally correlated. Older dams, which pose a greater safety risk, may be concentrated further upstream of state borders. Alternatively, states could favor dam removals that benefit native cold-water species. Colder waters tend to be located in higher elevations and lower stream orders, which correlate with upstream distance to state borders in our study area.

Regardless of the underlying mechanism, our results suggest a discrepancy between state and regional objectives. Inefficiencies from state-level environmental rulemaking have been uncovered in air quality (Gray and Shadbegian, 2004), water quality (Gray and Shadbegian, 2004; Lipscomb and Mobarak, 2017; Sigman, 2005, 2002), and water quantity (Olmstead and Sigman, 2015). Our findings provide initial evidence that environmental federalism contributes to inefficient production of ecological restoration outcomes. Coordinating barrier removals within and across river basins can improve the efficiency of river restoration efforts (Fitzpatrick and Neeson, 2018; Neeson et al., 2015; Roy et al., 2018). However, our results suggest that state-level decision making poses an institutional barrier to broader coordination. The implications extend to other spatially explicit conservation strategies looking to maximize biological gains by reducing habitat fragmentation, including terrestrial protected areas (Fremier et al., 2015). Operationalizing these types of strategies, therefore, may require the application of existing or novel incentive structures to align state and federal conservation objectives.

Finally, we note a methodological finding. Our estimation results were consistent between binary and continuous treatment variables and matching methods despite a relatively large difference in the number of observations (9,505 for binary matching and 11,150 for continuous matching). One study in the economics literature comparing results from continuous and binary matching (Cerulli et al., 2018) found significant loss of information from dichotomizing their treatment variable. In contrast, our goal with matching was to preprocess the data prior to model estimation (Ho et al., 2007). For data preprocessing applications like ours, the choice between matching methods appears to hold little importance.

8. Conclusion

In this paper we asked whether environmental federalism, or decentralized decision making, is a determinant of which dams are chosen for removal. Dam removals are an increasingly common approach to river restoration, owing to their effectiveness for restoring ecosystem function and recovering native fish populations (Bednarek, 2001). Coordinating removal decisions across larger geographies leads to more efficient outcomes (Roy et al., 2018), but states may have more inward-looking objectives.

We constructed a comprehensive and spatially explicit database of existing and removed dams in order to model the determinants of dam removal. To test for an environmental federalism story, we preprocessed our dataset using two matching strategies and employed an instrumental variable to identify the effect of state border

proximity on dam removal likelihood. Our results indicate dams located upstream of a state border are about 7 percentage points more likely to be removed than similar dams with no downstream border. Removal likelihood increases for dams located farther upstream of the border at a rate of between 5.8 and 7.0 percentage points per standard deviation. Our results are qualitatively the same across different models and matching strategies.

We consider our results evidence that state-level dam removal decisions diverge from larger scale prioritization or optimization efforts. However, the overall effect on social welfare depends on the extent to which benefits from dam removal are a global or local public good. The closer benefits are to a global (local) public good, the greater (smaller) the welfare loss from decentralized decisions. A nuanced analysis may uncover that states actually are best equipped to make socially optimal location decisions for restoration projects. Though beyond the scope of this paper, it is an important area for potential research in the area of environmental federalism in restoration and conservation more generally. We highlight additional research opportunities below.

As the number of dam removals expands over time and space, this will continue to be an appropriate setting to study the consequences of environmental federalism for efficient restoration and conservation objectives. More data can mean better identification. For instance, some dams in the central and western US may predate the drawing of state borders. For that subpopulation, downstream distance to the border is plausibly random, especially when the border is a straight line corresponding to a parallel or meridian. Another potential with expanded data is to

look for possible mitigating factors such as the strength of other interstate watershed coordination agreements. Such studies will serve a larger goal of designing frameworks to operationalize large scale, efficiency-based conservation and restoration strategies.

References

- American Rivers, 2018. American Rivers Dam Removal Database. figshare. Dataset.
- Auerbach, D., Deisenroth, D., McShane, R.R., McCluney, K.E., LeRoy Poff, N., 2014. Beyond the concrete: Accounting for ecosystem services from free-flowing rivers. *Ecosyst. Serv.* 10, 1–5.
<https://doi.org/10.1016/j.ecoser.2014.07.005>
- Bednarek, A.T., 2001. Undamming Rivers: A Review of the Ecological Impacts of Dam Removal. *Environ. Manage.* 27, 803–814.
<https://doi.org/10.1007/s002670010189>
- BenDor, T., Lester, T.W., Livengood, A., Davis, A., Yonavjak, L., 2015. Estimating the Size and Impact of the Ecological Restoration Economy. *PLoS ONE* 10.
<https://doi.org/10.1371/journal.pone.0128339>
- Berhane, G., Gilligan, D.O., Hoddinott, J., Kumar, N., Taffesse, A.S., 2014. Can Social Protection Work in Africa? The Impact of Ethiopia’s Productive Safety Net Programme. *Econ. Dev. Cult. Change* 63, 1–26.
- Besley, T., Coate, S., 2003. Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach. *J. Public Econ.* 87, 2611–2637.
- Birthal, P.S., Roy, D., Negi, D.S., 2015. Assessing the Impact of Crop Diversification on Farm Poverty in India. *World Dev.* 72, 70–92.
<https://doi.org/10.1016/j.worlddev.2015.02.015>
- Bloch, F., Zenginobuz, U., 2007. The effect of spillovers on the provision of local public goods. *Rev. Econ. Des. Heidelb.* 11, 199.
<http://dx.doi.org.uri.idm.oclc.org/10.1007/s10058-006-0016-x>
- Braid, R.M., 2010. Provision of a pure local public good in a spatial model with many jurisdictions. *J. Public Econ.* 94, 890–897.
<https://doi.org/10.1016/j.jpubeco.2010.07.008>
- Cerulli, G., Potì, B., Spallone, R., 2018. The impact of fiscal relief on multinationals business R&D investments: a cross-country analysis. *Econ. Polit.* 35, 649–675.
<https://doi.org/10.1007/s40888-018-0123-5>
- Costanza, R., 2008. Ecosystem services: Multiple classification systems are needed. *Biol. Conserv.* 141, 350–352. <https://doi.org/10.1016/j.biocon.2007.12.020>
- Diessner, N.L., Ashcraft, C.M., Mo, W., Song, C., 2020. Pearl River Negotiation Simulation: Negotiating the Future of Dams. University of New Hampshire, Durham.

- Ding, L., Chen, L., Ding, C., Tao, J., 2019. Global Trends in Dam Removal and Related Research: A Systematic Review Based on Associated Datasets and Bibliometric Analysis. *Chin. Geogr. Sci.* 29, 1–12. <https://doi.org/10.1007/s11769-018-1009-8>
- Doyle, M., Harbor, J., Stanley, E., 2003. Toward Policies and Decision-Making for Dam Removal. *Environ. Manage.* 31, 0453–0465. <https://doi.org/10.1007/s00267-002-2819-z>
- Fitzpatrick, K.B., Neeson, T.M., 2018. Aligning dam removals and road culvert upgrades boosts conservation return-on-investment. *Ecol. Model.* 368, 198–204. <https://doi.org/10.1016/j.ecolmodel.2017.11.018>
- Fong, C., Hazlett, C., Imai, K., 2018. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. *Ann. Appl. Stat.* 12, 156–177. <https://doi.org/10.1214/17-AOAS1101>
- Fong, C., Ratkovic, M., Hazlett, C., Imai, K., 2017. CBPS: R package for covariate balancing propensity score. Available at the Comprehensive R Archive Network (CRAN): <https://CRAN.R-project.org/package=CBPS>.
- Fox, C.A., Magilligan, F.J., Sneddon, C.S., 2016. “You kill the dam, you are killing a part of me”: Dam removal and the environmental politics of river restoration. *Geoforum* 70, 93–104. <https://doi.org/10.1016/j.geoforum.2016.02.013>
- Fox, E., Klein, S., Roy, S.G., Raffier, K., 2019. Participatory multi-criteria decision analysis. *Future of Dams Project Research Brief*.
- Fremier, A.K., Kiparsky, M., Gmur, S., Aycrigg, J., Craig, R.K., Svancara, L.K., Goble, D.D., Cosens, B., Davis, F.W., Scott, J.M., 2015. A riparian conservation network for ecological resilience. *Biol. Conserv.* 191, 29–37. <https://doi.org/10.1016/j.biocon.2015.06.029>
- Gold, A.J., Addy, K., Morrison, A., Simpson, M., 2016. Will Dam Removal Increase Nitrogen Flux to Estuaries? *Water* 8, 522. <https://doi.org/10.3390/w8110522>
- Graber, Singler, McClain, Thomas-Blate, 2015. Removing Small Dams: A Basic Guide for Project Managers. *American Rivers*.
- Grabowski, Z.J., Chang, H., Granek, E.L., 2018. Fracturing dams, fractured data: Empirical trends and characteristics of existing and removed dams in the United States. *River Res. Appl.* 0. <https://doi.org/10.1002/rra.3283>
- Graf, W.L., 1999. Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. *Water Resour. Res.* 35, 1305–1311. <https://doi.org/10.1029/1999WR900016>

- Gray, W.B., Shadbegian, R.J., 2004. ‘Optimal’ pollution abatement—whose benefits matter, and how much? *J. Environ. Econ. Manag.*, Including Special Symposium Section from the National Bureau of Economic Research Conference on Advances in Empirical Environmental Policy Research 47, 510–534. <https://doi.org/10.1016/j.jeem.2003.01.001>
- Ho, D.E., Imai, K., King, G., Stuart, E.A., 2011. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J. Stat. Softw.* 42, 1–28.
- Ho, D.E., Imai, K., King, G., Stuart, E.A., 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Polit. Anal.* 15, 199–236. <https://doi.org/10.1093/pan/mpi013>
- Imai, K., King, G., Stuart, E.A., 2008. Misunderstandings between experimentalists and observationalists about causal inference. *J. R. Stat. Soc. Ser. A Stat. Soc.* 171, 481–502. <https://doi.org/10.1111/j.1467-985X.2007.00527.x>
- Imai, K., Ratkovic, M., 2014. Covariate balancing propensity score. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 76, 243–263. <https://doi.org/10.1111/rssb.12027>
- Imbens, G.W., 2000. The role of the propensity score in estimating dose-response functions. *Biometrika* 87, 706–710. <https://doi.org/10.1093/biomet/87.3.706>
- Ishiyama, N., Ryo, M., Kataoka, T., Nagayama, S., Sueyoshi, M., Terui, A., Mori, T., Akasaka, T., Nakamura, F., 2018. Predicting the ecological impacts of large-dam removals on a river network based on habitat-network structure and flow regimes. *Conserv. Biol.* 32, 1403–1413. <https://doi.org/10.1111/cobi.13137>
- Kraft, M., Rosenberg, D.E., Null, S.E., 2019. Prioritizing Stream Barrier Removal to Maximize Connected Aquatic Habitat and Minimize Water Scarcity. *JAWRA J. Am. Water Resour. Assoc.* 55, 382–400. <https://doi.org/10.1111/1752-1688.12718>
- Lipscomb, M., Mobarak, A.M., 2017. Decentralization and Pollution Spillovers: Evidence from the Re-drawing of County Borders in Brazil. *Rev. Econ. Stud.* 84, 464–502. <https://doi.org/10.1093/restud/rdw023>
- Magilligan, F.J., Graber, B.E., Nislow, K.H., Chipman, J.W., Sneddon, C.S., Fox, C.A., 2016. River restoration by dam removal: Enhancing connectivity at watershed scales. *Elem. Sci. Anthr.* 4, 000108. <https://doi.org/10.12952/journal.elementa.000108>
- Magrini, E., Montalbano, P., Nenci, S., 2017. Are EU trade preferences really effective? An impact evaluation assessment of the Southern Mediterranean Countries’ case. *Int. Rev. Appl. Econ.* 31, 126–144. <https://doi.org/10.1080/02692171.2016.1222355>

- Martin, E., Apse, C., 2011. Northeast Aquatic Connectivity: An Assessment of Dams on Northeastern Rivers. The Nature Conservancy, Eastern Freshwater Program.
- Martin, E.H., Levine, J., 2017. Northeast Aquatic Connectivity Assessment Project - Version 2.0: Assessing the ecological impact of barriers on Northeastern rivers. The Nature Conservancy, Brunswick, Maine.
- Neeson, T.M., Ferris, M.C., Diebel, M.W., Doran, P.J., O’Hanley, J.R., McIntyre, P.B., 2015. Enhancing ecosystem restoration efficiency through spatial and temporal coordination. *Proc. Natl. Acad. Sci.* 112, 6236–6241. <https://doi.org/10.1073/pnas.1423812112>
- Oates, W.E., 1972. *Fiscal Federalism*. Harcourt Brace, New York.
- Olmstead, S.M., Sigman, H., 2015. Damming the Commons: An Empirical Analysis of International Cooperation and Conflict in Dam Location. *J. Assoc. Environ. Resour. Econ.* 2, 497–526. <https://doi.org/10.1086/683205>
- Poff, N.L., Hart, D.D., 2002. How Dams Vary and Why It Matters for the Emerging Science of Dam Removal. *BioScience* 52, 659–668. [https://doi.org/10.1641/0006-3568\(2002\)052\[0659:HDVAWI\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0659:HDVAWI]2.0.CO;2)
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Roy, S.G., Uchida, E., Souza, S.P. de, Blachly, B., Fox, E., Gardner, K., Gold, A.J., Jansujwicz, J., Klein, S., McGreavy, B., Mo, W., Smith, S.M.C., Vogler, E., Wilson, K., Zydlewski, J., Hart, D., 2018. A multiscale approach to balance trade-offs among dam infrastructure, river restoration, and cost. *Proc. Natl. Acad. Sci.* 115, 12069–12074. <https://doi.org/10.1073/pnas.1807437115>
- Sigman, H., 2005. Transboundary spillovers and decentralization of environmental policies. *J. Environ. Econ. Manag.* 50, 82–101. <https://doi.org/10.1016/j.jeem.2004.10.001>
- Sigman, H., 2002. International spillovers and water quality in rivers: Do countries free ride? *Am. Econ. Rev.* 92, 1152–1159.
- Smith, S.V., Renwick, W.H., Bartley, J.D., Buddemeier, R.W., 2002. Distribution and significance of small, artificial water bodies across the United States landscape. *Sci. Total Environ.* 299, 21–36. [https://doi.org/10.1016/S0048-9697\(02\)00222-X](https://doi.org/10.1016/S0048-9697(02)00222-X)

- Terza, J.V., Basu, A., Rathouz, P.J., 2008. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *J. Health Econ.* 27, 531–543. <https://doi.org/10.1016/j.jhealeco.2007.09.009>
- U.S. Census Bureau, 2017. American Community Survey, 2017 5-Year Public Use Microdata.
- Veall, M.R., Zimmermann, K.F., 1996. Pseudo-R² Measures for Some Common Limited Dependent Variable Models. *J. Econ. Surv.* 10, 241–259. <https://doi.org/10.1111/j.1467-6419.1996.tb00013.x>
- Weir, M.J., Ashcraft, C.M., Diessner, N.L., McGreavy, B., Vogler, E., Guilfoos, T., 2020. Language effects on bargaining. *PLOS ONE* 15, e0229501. <https://doi.org/10.1371/journal.pone.0229501>

Figure 1. Dam removal frequency by year (1990-2017)
Note: N=396

Figure 2. Number of existing and removed dams by state as of 2017
Notes: N=11,601. States are ordered by frequency of removals

Figure 3. Geospatial distribution of treatment and control groups
Note: N=11,601.

Figure 4. Comparison of densities for main confounder before and after matching
 Notes: Downstream distance to river mouth (km) is along x-axis. Observations from control group (no state border downstream) are weighted according to their frequency as matches.

Table 1. Result of t-tests for pre-matching difference in means by outcome variable and treatment

Variable	Existing	Removed	t-stat	Upstream of Border	Not Upstream of Border	t-stat
Proportion removed	0	1		0.043	0.022	6.51
Proportion with a state border on downstream flowpath	0.571	0.727	6.84	1	0	
Downstream distance to state border, if one exists (km)	106.9	118.4	1.59	107.4		
Physical controls						
Downstream distance to river mouth (km)	193.6	198.2	0.69	276.2	81.7	82.20
Number of downstream barriers	7.0	6.0	3.96	8.72	4.48	44.64
Downstream passability product	0.1111	0.1212	0.61	0.0620	0.1786	18.78
Total functional network (km)	361.4	544.8	2.66	504.8	181.3	18.46
River size class	1.79	2.38	11.16	1.84	1.77	4.10
Slope	0.013	0.0107	3.00	0.0155	0.0095	20.62
Proportion on a state border	0.004	0.008	0.81			
Ecological controls						
Landcover in contributing watershed						
% Impervious	4.62	4.60	0.06	3.12	6.66	23.87
% Agriculture	11.46	18.8	7.79	13.28	9.59	12.84
% Natural	70.84	63.02	5.90	73.29	66.88	14.37
Proportion falling on conserved land	0.1695	0.1616	0.42	0.1728	0.1644	1.20
Cumulative disturbance index	2.11	2.01	1.33	2.27	1.88	13.77
Number of anadromous species downstream	0.74	1.11	3.96	0.57	1.01	14.32
Species richness (within HUC-8)	35.5	37.2	3.55	36.5	34.3	11.29
Demographic controls (Census tract-level)						
Population density (per sq km)	0.78	1.09	2.43	0.57	1.08	12.97
Per capita income (\$1,000s)	38.60	33.93	7.23	35.03	43.07	25.74
% Collecting Social Security	35.61	35.77	0.38	36.55	34.36	14.12
Proportion of sample	0.966	0.034		0.5761	0.4239	

Notes: N=11,601. Bold indicates p-value ≤ 0.10 .

Table 2. Result of t-tests for difference in means on matched subsamples

Variable	Binary Matching		Continuous Matching	
	Upstream of Border	Not Upstream of Border	Upstream of Border	Not Upstream of Border
Physical controls				
Downstream distance to river mouth (km)	276.2	279.2	0.78	41.21
Number of downstream barriers	8.5	7.3	1.75	20.23
Downstream passability product	0.062	0.043	3.98	9.28
Total functional network (km)	503.4	568.0	2.96	11.26
River size class	1.84	1.86	1.02	1.07
Slope	0.016	0.014	4.9	16.39
Ecological controls				
Landcover in contributing watershed				
% Impervious	3.1	3.0	0.59	4.9
% Agriculture	13.3	13.8	1.46	11.3
% Natural	73.3	73.0	0.76	70.5
Proportion falling on conserved land	0.173	0.175	0.30	0.46
Cumulative disturbance index	2.27	2.04	6.75	9.66
Number of anadromous species downstream	0.56	0.56	0.27	4.26
Species richness (within HUC-8)	36.50	39.89	13.14	36.15
Demographic controls (Census tract-level)				
Population density (per sq km)	0.57	0.60	0.36	0.79
Per capita income (\$1,000s)	35.04	33.99	3.61	13.91
% Collecting Social Security	36.6	37.2	3.4	36.38
N	6681	2824		4479
Proportion of sample (unweighted)		0.297		0.401

Notes: All variables were used to obtain optimized propensity scores based on binary (presence/absence of downstream border) and continuous (downstream distance to state border) definitions of treatment. Propensity scores are used in subsequent 3:1 nearest neighbor matching. Treated observations are constant across samples. Control observations are weighted according to their frequency as matches. Bold indicates p-value ≤ 0.10 .

Table 3. Marginal effects from logistic regression

	(1)	(2)	(3)	(4)	(5)
Upstream of state border (D=1)	0.0287*** (0.0057)	0.0415*** (0.0071)	0.0175* (0.0090)	0.0378*** (0.0097)	0.0275*** (0.0097)
Distance from state border (10km)	0.0003* (0.0002)	0.0009*** (0.0002)	0.0008*** (0.0003)	0.0006** (0.0003)	0.0009*** (0.0003)
Control variables		Yes	Yes	Yes	Yes
State fixed effects			Yes	Yes	Yes
Matching to preprocess data				Yes	Yes
Treatment variable				Binary	Continuous
N	11,601	11,601	11,601	9,505	11,160
Pseudo R-squared	0.0355	0.1797	0.4501	0.5110	0.4624

Notes: Dependent variable is status of the dam (0 = existing, 1 = removed). *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Table 4. IV Results

First Stage	(1)	(2)	(3)
Percent commuting out of state	-0.1569*** (0.0073)	-0.1906*** (0.0091)	-0.1746*** (0.0078)
F-stat	450.1	834.7	626.7
Second Stage			
Upstream of state border (D=1)	0.0609*** (0.0153)	0.0731*** (0.0174)	0.0706*** (0.0151)
Distance from state border (10km)	0.0056*** (0.0013)	0.0053*** (0.0012)	0.0061*** (0.0012)
Pseudo R-squared	0.4622	0.5161	0.4748
N	11,601	9,505	11,160
Matching to preprocess data		Yes	Yes
Treatment variable		Binary	Continuous

Notes: First stage is OLS and dependent variable is downstream distance to state border. Second stage are marginal effects from a logistic regression where the dependent variable is status of the dam (0 = existing, 1 = removed) and residuals from first stage are included as a regressor. Both stages include full suite of controls including state fixed effects. Second stage standard errors are bootstrapped using 1,000 replications. *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Integrating public preferences with biophysical production possibilities:
an application to ecosystem services from dam removal

(To be submitted to Ecology and Society)

Ben Blachly¹, Emi Uchida, Samuel G. Roy

¹ Ben Blachly is corresponding author and is a PhD candidate in the Department of Environmental and Natural Resource Economics, University of Rhode Island (ben_blachly@uri.edu). Emi Uchida is a Professor in the Department of Environmental and Natural Resource Economics, University of Rhode Island. Samuel G. Roy is an Assistant Research Professor at the Senator George J. Mitchell Center for Sustainability Solutions, University of Maine. Support for this project is provided by the National Science Foundation's Research Infrastructure Improvement NSF #IAA-1539071.

ABSTRACT

Increasing the provision of one ecosystem service often comes only at the expense of others. Effective management requires understanding the biophysical relationships governing these tradeoffs, as well as stakeholder preferences for the tradeoffs. In this paper we apply microeconomic principles of production and consumption to empirically identify efficient and socially desirable ecosystem service outcomes. We use a production possibilities model to quantify the tradeoffs among four ecosystem services from dam removal in Maine: hydropower, lake shoreline, herring, and Atlantic salmon. Then we conduct a choice experiment survey to understand public preferences for the services, using latent class analysis to group respondents with common preferences. Finally, we estimate indifference curves and combine them with the production possibilities model to identify preferred ecosystem service outcomes by group. Results indicate that stakeholders with apparently diverging values can sometimes agree on preferred outcomes when we account for the productive capacity of the ecosystem. Additionally, we find that the common practice of reducing multidimensional ecosystem service tradeoff analyses to two dimensions can result in lost welfare. Generally, our empirical application demonstrates the practical usability of this framework for incorporating diverse stakeholder preferences into ecosystem management planning.

1. Introduction

Ecosystem services are the benefits that humans derive from the environment. Because of the biophysical relationships governing their production, managing ecosystem services often implies increasing some at the expense of others (Bennett et al., 2009). Significant progress has been made developing models to quantify these tradeoffs using the economic concept of a production possibilities frontier (PPF) (e.g., Ager et al., 2017; Roy et al., 2018; White et al., 2012), but the biophysical models do not tell us which outcomes are most desirable. Methodologies for incorporating stakeholder preferences are needed to transform the models into viable tools for decision making, especially when stakeholder groups have different values (Cord et al., 2017). One promising framework is firmly grounded in welfare economics, using the concept of indifference curves to identify utility-maximizing combinations of ecosystem services from a production possibilities model (Cavender-Bares et al., 2015). While the general economic theory extends to higher dimensions, it has been suggested that reducing analyses to two dimensions is desirable to utilize the framework for decision making (King et al., 2015).

This paper asks how the PPF – indifference curve framework can be implemented empirically to facilitate ecosystem service management decisions, and it studies the implications of simplifying multidimensional tradeoffs to two dimensions. In empirical simulations of ecosystem service production possibilities, it is common to bundle correlated services into two groups or to consider a single, representative two-way tradeoff (e.g., Nalle et al., 2004; Polasky et al., 2008). Doing so facilitates

visualization, comprehension, and computation. In a decision-making framework intended to integrate variation in stakeholder preferences, comprehension through visualization is especially important. However, there is no guarantee that preferences and production will be correlated along the same lines. If this is the case, the socially desirable level of one service will depend on which competing service it is paired against. As a result, welfare could suffer, stakeholder conflicts could arise unnecessarily, and opportunities for consensus could be missed.

This paper makes a couple of important contributions. First, it is the first empirical application combining a production possibilities model with utility-theoretic indifference curves to identify socially preferred ecosystem service outcomes. Cavender-Bares et al. (2015) proposed applying the microeconomic framework to navigate ecosystem tradeoff decisions. In a follow up paper, King et al. (2015) use hypothetical indifference curves and production models to facilitate stakeholder discussions in a workshop setting. We are not aware of any empirical examples of the explicit theoretical framework that adhere to its microeconomic foundations. Thus, we contribute by demonstrating a practical method for implementing the framework using a choice experiment to estimate indifference curves.

Second, this research is unique in that the choice experiment is designed in conjunction with the production possibilities model. The consistency allows us to make meaningful comparisons of the insights gained from analyzing the production possibilities model, the choice experiment, and the combined model. Hattam et al. (2015) compare results of an ecological assessment with a choice experiment,

concluding that the two are complementary. Our work goes one step further by investigating additional insights gained from combining the two in a single model.

We set our study in the context of watershed-scale decisions about dam removal in the state of Maine, USA. The dichotomy presented by dams and their removal provides an ideal application for analyzing tradeoffs between services. We use focus groups to identify four relevant ecosystem services: hydropower production, lake shoreline, spawning habitat for herring, and spawning habitat for Atlantic salmon. We utilize production functions for each service to create two-dimensional PPFs for pairwise analysis of each dam-dependent service (hydropower and lakes) against each that benefits from dam removal (herring and salmon populations). Each PPF identifies Pareto efficient outcomes (i.e., combinations where production of one service cannot be increased without decreasing production of the other). We compare the two-dimensional tradeoff results to those obtained from a four-dimensional PPF, which identifies four-way Pareto efficient outcomes by accounting for the production relationships between all four services simultaneously.

Next we administer a choice experiment survey to understand public preferences for the same four ecosystem services. We use the PPF results to identify feasible levels of each service, which we use to create an efficient experimental design. We analyze the results using a latent class model to identify groups holding similar preferences. We estimate each group's willingness to pay (WTP) for each service, as well as the rate at which they are willing to trade one service for another (i.e., the marginal rates of substitution). We use these marginal rates of substitution to construct indifference curves, which we combine with the PPF to identify which

Pareto efficient outcomes each group finds most desirable. Finally, we compare the results derived from two-dimensional analyses to the results from four-dimensional analysis in terms of a monetary welfare measure.

We find significant advantages for decision making from analyzing the PPF and choice experiment results jointly. For example, our choice experiment data indicates significant variation across stakeholder groups in WTP for the ecosystem services. We find, however, that variations in WTP do not necessarily result in large disagreements about desired outcomes. The magnitude of disagreement depends on the shape of the PPF, and seemingly large disagreements may actually be noncontentious. Additionally, in contrast to WTP analysis, utilizing indifference curves in conjunction with PPFs enables us to incorporate preferences for stakeholders who refuse to value ecosystem services monetarily.

We also find that simplifying multidimensional problems to two dimensions can have undesirable consequences. Bundling or proxying sets of correlated ecosystem services into two groups can be a convenient way to facilitate visualization, comprehension, and computation of tradeoffs (King et al., 2015). However, our two-way results indicate that the preferred level of a service often depends on which service it is paired against. Further, we find welfare losses associated with identifying preferred ecosystem service outcomes in two dimensions compared to a four-dimensional analysis.

2. Concepts

Grounded in economic production theory, a production possibilities frontier (PPF) identifies the maximum amount of a good or service that can be produced holding production of other goods and services constant. Combinations of goods and services that define the PPF are efficient in the sense that productive capacity is fully utilized. When capacity is fully utilized, increasing production of one good or service can come only at the expense of others. In this way the PPF allows us to quantify and visualize tradeoffs, particularly in simple, two-dimensional cases along the familiar x-y plane.

A PPF can easily be applied to the production of ecosystem services, and it can be a helpful tool for decision makers (Vallet et al., 2018). In this context, a PPF identifies combinations of services that fully utilize the productive capacity of an ecosystem. Some researchers reduce higher dimension problems to two dimensions, presumably to ease computation and facilitate visualization. One approach is to group correlated ecosystem services (King et al., 2015). For example, Polasky et al. (2008) group vertebrate species and industrial activities in order to analyze the two-way tradeoff between biodiversity and economic returns. Another approach is to focus on the main tradeoff between two primary services, which become de facto proxies for other, correlated services. Mastrangelo and Laterra (2015) study the tradeoffs associated with deforestation in Argentina by focusing on cattle production and avian biodiversity. Simplifications like these accurately describe the biophysical tradeoffs

between the two ecosystem services, but problems could arise as we incorporate stakeholder preferences, which may include a broader range of services.

Integrating preferences into PPF models can transform the biophysical model into a useful decision-support tool (Cord et al., 2017). Theoretically, PPF analysis can be combined with stakeholder preferences represented by indifference curves to identify preferred outcomes (Cavender-Bares et al., 2015; King et al., 2015). The points along an indifference curve represent combinations of ecosystem services that deliver a fixed level of utility, or well-being. The slope indicates how much of one service an individual is willing to trade off to gain additional units of another service without sacrificing utility. Thus, it represents preferences for tradeoffs between services. To theoretically maximize utility in a two-dimensional example, an individual wants to be on an indifference curve as far from the origin as possible (i.e., where she is receiving more of both services). She is bound, however, by the biophysical constraints of the ecosystem, represented by the PPF. The utility maximizing solution, therefore, occurs at the point of tangency between PPF and indifference curve. At this point the slopes are equated—stakeholder willingness to tradeoff matches the ecosystem’s ability to generate tradeoffs.

Despite its theoretical attractiveness, applications of the PPF – indifference curve framework are relatively rare. King et al. (2015) utilize the framework in workshops, creating hypothetical PPFs and indifference curves to identify potential solutions when there is variation in stakeholder preferences. In a similar application, Stosch et al. (2019) ask stakeholders their beliefs about the shape of a PPF, and how they perceive others’ preferences for the tradeoffs in order to facilitate discussion. In

the sole empirical example that we could find, Mastrangelo and Laterra (2015) construct indifference curves using a Likert scale measure of landholders' intent to engage in conservation practices, which they combine with an empirical PPF to identify preferred outcomes. We contribute to this general literature by incorporating an established technique for estimating indifference curves that is consistent with the utility-theoretic underpinning of the PPF – indifference curve framework.

Discrete choice experiments, a common nonmarket valuation method employed by environmental economists, allow straightforward estimation of indifference curve slopes. In a choice experiment, respondents face a set of options described by a series of attributes. In an ecosystem services context, each option may represent a management plan described by the levels of services it is expected to deliver. The analyst assumes the respondent chooses the option that maximizes her utility, which depends in turn on the attribute levels.¹ Researchers can then estimate the relative contribution of each attribute to respondent utility, or the marginal utility of each attribute. The ratio of marginal utilities from two attributes indicates the rate at which respondents are willing to trade one attribute for another while maintaining a constant level of utility. This marginal rate of substitution is identical to the slope of an indifference curve. Choice experiments, therefore, are a natural method for implementing the PPF – indifference curve framework empirically.

¹ More formally, choice modelling is based on random utility theory, which assumes that only a portion of utility can be described by the attribute levels. The additional portion is considered unobservable, or random, following a specified distribution. See Train (2009) for a complete technical treatment.

3. Study context

We set our study in the context of dam removal decision making, a ripe application for the PPF – indifference curve framework. Dams are a ubiquitous part of the modern American landscape, numbering over two million nationwide (Smith et al., 2002). Growing appreciation for the ecosystem services provided by freely flowing rivers (Auerbach et al., 2014), though, has led to an increase in purposeful dam removals. To date, over 1,400 dams have been removed in the US, with the vast majority occurring in the past two decades (American Rivers, 2018). Currently, dams tend to be selected for removal opportunistically (Magilligan et al., 2016), and decisions can be governed by emotion over science (Fox et al., 2016). However, the potential efficiency gains from coordinating dam removal decisions across large scales are well understood (Neeson et al., 2015; Roy et al., 2018). Decision support tools are needed that can mitigate variation in stakeholder preferences to realize efficiency gains from coordination.

We focus specifically on dam removal in the Penobscot Watershed in Maine, USA (Figure 1). The watershed boasts a drainage area of over 8,500 square miles, roughly one third of the state, and is home to 112 dams spread across 7,000 river miles. The large number of dams translates to many possible dam removal scenarios, which allows us to construct a PPF with meaningful tradeoffs. Additionally, the Penobscot is an attractive candidate for watershed-scale policy coordination. Despite its size it is located entirely within a single state jurisdiction. Also, the Penobscot has

some, albeit limited, history of coordinating removal decisions across multiple dams (Opperman et al., 2011).

We held focus groups in the study area to identify four ecosystem services on which to focus. Two rely on the presence of dams: hydropower and dammed lakes, and two benefit from dam removal: river herring and Atlantic salmon populations. Hydropower in Maine is significant, accounting for nearly one third of net electricity production in the state (United States EIA, 2019). Within the Penobscot Watershed, 17 dams are licensed to generate hydropower for a total installed capacity of 167,469 kW, enough to supply about 155,000 homes with relatively clean energy. Dams also support and stabilize the level of lakes, which have recreational and aesthetic value. There are approximately 3,020 miles of lake shoreline in the Penobscot Watershed, of which roughly 500 miles depend on the presence of dams.

The main beneficiaries of dam removal are fish, particularly river herring and Atlantic salmon. Both are anadromous species, meaning they spawn in freshwater river systems, migrate to the ocean as juveniles, and return to freshwater to spawn subsequent generations. Because dams physically block up and downstream migration, they are a known cause of declining fisheries (Nations, 2001) and they have contributed to the current listing of Atlantic salmon as an endemically endangered species (Limburg and Waldman, 2009). Atlantic salmon hold intrinsic value, while river herring hold joint importance to the ecosystem as a keystone species and to the Maine economy as a source of bait for the state's lobster fishery. The current configuration of dams in the Penobscot Watershed provides accessible habitat to

support about 5,000 Atlantic salmon and 260,000 river herring (23% and 2% of the respective historic populations).

4. Methods

4.1 Production possibilities frontier

To generate the PPF, we used the New England Dams Database (<https://ddc-nedams.sr.unh.edu/>) combined with hydropower capacity licensing data from the Federal Energy Regulatory Commission (FERC, 2019), hydrologic data (USGS, 2017), and fish passage and habitat suitability models (Roy et al., 2018). There are 112 dams in the Penobscot Watershed. Of these dams, 17 are licensed to generate hydropower, 58 have a lake or pond reservoir, and 62 block historic upstream habitat for Atlantic salmon. We followed the methods in Roy et al. (2018) to measure upper and lower limits of attribute values based on hypothetical scenarios in which all dams in the basin are either kept in current condition or completely removed. While hydropower capacity is simply calculated as the cumulative sum of capacity across all generating dams, Atlantic salmon carrying capacity requires a spatially explicit ecological model that we describe below.

Population estimates for Atlantic salmon are based on available habitat extent, quality, and accessibility data (Roy et al., 2018)

$$S = c \sum_{i \in n_d} \left[h_i \prod_{j \in n_{di}} (p_j) \right] \quad (1)$$

where S is annual Atlantic salmon spawning population capacity (number of fish); n_d is the set of all dams, indexed by i ; n_{di} is the set of all dams downstream from and including dam i , indexed by j ; h_i is the functional habitat accessible above dam i (km^2); p_j is the upstream and downstream product of probability of passage through downriver dam j ; and c is the annual population carrying capacity (number of fish $\text{m}^{-2} \text{a}^{-1}$). We determine the spatial distribution of functional habitat h_i based on physical surveys and historic accounts (Houston et al., 2007; Martin and Apse, 2011), and estimate functional habitat quality based on environmental (Olivero and Anderson, 2008) and hydrologic (USGS, 2017) model data and suitability requirements (Armstrong et al., 2003; NOAA, 2009)

$$h_i = \sum_{r \in g_i} [a_r t_r v_r \vartheta_r \cdot 0.8] \quad (2)$$

where g_i is the set of river reaches immediately above dam i , indexed by r , a_r is the mean annual wetted area within reach r (m^2), t_r is the mean annual temperature factor in reach r , v_r is the mean annual velocity factor in reach r , ϑ_r is a binary value identifying if the river reach r is historically accessible to Atlantic salmon, and 0.8 is included to assume that 20% of available surface area exists along the river banks and is unsuitable for habitat use due to natural river height fluctuations. Locations with seasonal drying or tidal effects are excluded from analysis to reflect species' long-term intolerance of these conditions for spawning and rearing (Armstrong et al., 2003; NOAA, 2009).

4.2 Choice experiment survey and modelling

We administered a choice experiment survey to residents of the Penobscot Watershed to elicit preferences for the ecosystem service tradeoffs from dam removal. The survey consisted of four parts. First, we asked respondents about their perceptions and usage of nearby rivers, and their prior knowledge of dam removal efforts. Next was an informational section that described the current situation in terms of each ecosystem service, and how the levels could be different in the future based on alternative dam removal plans. This was immediately followed by the choice experiment. We used feasible ranges for each ecosystem service from the PPF analysis to create an efficient experimental design (Rose and Bliemer, 2009). Each respondent faced six choices (see Figure 2 for an example). In the final section we asked a series of demographic questions and questions to elicit environmental orientation (Dunlap et al., 2000). Based on reports of low broadband penetration in the study area, we opted for a postal mail survey utilizing five points of contact (Dillman et al., 2014). Data collection took place between September and November of 2018. We received 204 responses with all six choice questions completed for a response rate of about 31%.

Choice experiment modelling is most commonly based on random utility theory (Train, 2009). We assume the respondent chooses the available option that maximizes her utility, which can be decomposed into observable and unobservable, or random, components. The most basic model is a multinomial logit (MNL) which assumes that the random components are independent and identically distributed and

follow an extreme value type I (Gumbel) distribution. Next, we specify a functional form for the observable portion of utility, v .

$$v_i = \beta_{SQ} + \beta_{salm}x_{salm_i} + \beta_{hyd}x_{hyd_i} + \beta_{her}x_{her_i} + \beta_{lake}x_{lake_i} + \beta_{cost}x_{cost_i} \quad (3)$$

Equation (3) says that v_i , the observable portion of utility from alternative i , is a linear combination of the level of Atlantic salmon (x_{salm_i}), hydropower (x_{hyd_i}), river herring (x_{her_i}), lake shoreline (x_{lake_i}), and the cost (x_{cost_i}) of that alternative. Each β_k is a parameter to be estimated, representing the marginal utility (i.e., the relative contribution) of each attribute k . The marginal rate of substitution between two attributes, x_k and x_j , is given by the ratio of marginal utilities, β_j/β_k . This tells us the rate at which respondents are willing to tradeoff one service for another without compromising utility, and is, by definition, the slope of an indifference curve. In nonmarket valuation applications, the researcher obtains an estimate of marginal WTP by finding the marginal rate of substitution between an attribute and the marginal utility of income, β_{cost} . Similarly, following Hanemann (1984), the total change in welfare associated with a new dam removal plan can be estimated as follows:

$$CS = -\frac{1}{\beta_{cost}} [\ln \sum_n \exp v_n^1 - \ln \sum_n \exp v_n^0] \quad (4)$$

where CS is the compensating surplus welfare measure and v_n^0 and v_n^1 are the n th residents' utility before and after the change, respectively.

A particularly useful application of the PPF – indifference curve framework is to identify areas of agreement and disagreement between groups. The basic MNL model, though, assumes preferences are fixed across respondents. By relaxing a few of the assumptions, we can obtain a latent class model (LCM). The LCM assumes that

each respondent belongs to a group which shares common preferences. The researcher can vary the number of groups, ultimately choosing based on model fit (Hensher et al., 2015) and the meaningfulness of parameter estimates (Scarpa and Thiene, 2005). Class membership is not observed by the researcher (i.e., classes are latent) but it can be predicted *ex post*, allowing us to characterize each class based on the demographic and attitudinal items presented in the survey. We use latent class analysis to create indifference curves for each class, representing groups of stakeholders with similar preferences.

4.3 Combining the results to identify preferred outcomes

Theoretically, the utility-maximizing combination of services is defined by the point of tangency between the PPF and indifference curve. In an empirical application, however, equating the slopes is impractical since the PPF is unlikely to be smooth. Further, it becomes visually and computationally intractable in higher dimensions. It is straightforward, though, to obtain the utility provided by each point along the PPF using Equation 3 once it has been parameterized. From there we simply select the point (i.e., the combination of services) that maximizes Equation 3 (the utility function). For our two-way analyses, visual depiction of the tangency condition could reveal insights that prove useful for facilitating discussions and negotiations. We produce these two-dimensional visualizations by plotting a line segment, with slope equal to the marginal rate of substitution, that intersects the PPF at the utility maximizing point.

5. Results and discussion

5.1 Production possibilities frontiers in two and four dimensions

The correlation between services from dams and between services from dam removal are apparent in the shape of the two-dimensional PPFs (Figure 3). Qualitative analysis using any of the two-way tradeoffs is similar. For all four PPFs, the current configuration of dams is represented by the upper leftmost endpoint, where hydropower and lake shoreline are maximized while both fish species are minimized. The results indicate that we are underutilizing the capacity of the Penobscot Watershed to produce services. Each graph features a flat segment at the top left, indicating the ability to increase either fish population without sacrificing hydropower or lake shoreline. In addition, all four PPFs are concave to the origin, representing the fact that the cost of restoring fish is increasing in terms of either hydropower or lake shoreline. Quantitative analysis, however, uncovers important differences.

Quantitatively, the two-dimensional results could be misleading if analyzed individually. For example, the lake shoreline results indicate that we can recover herring and Atlantic salmon to about 75% and 80% of their respective maximums with virtually no loss to lakes. However, it is unclear from that analysis how hydropower capacity would be affected. If we instead consider the tradeoff between both fish species and hydropower, restoring herring or Atlantic salmon to the same levels requires a 13% or 32% reduction in hydropower, respectively. We have no way to

know, though, how lake shoreline would be affected by the reduction in hydropower. To fully understand the interactions requires analyzing all four ecosystem services jointly.

The four-dimensional analysis yields a surface defined by over 15,000 unique combinations of dam removals. As in two-dimensions, each combination represents a Pareto-efficient outcome—a situation where an increase to one service can only come at the expense of at least one other. While a four-dimensional plot is feasible, its usefulness for understanding tradeoffs and informing decision making is questionable. We instead plot in two-dimensions (Figure 4), arbitrarily choosing hydropower-salmon space, only to note that each two-dimensional PPF (Figure 3) is a subset of the four-dimensional results. As an alternative to visualization, we can study the efficient outcomes in tabular form. We can learn, for example, that recovering herring to 75% of maximum with virtually no loss to lake shoreline (an apparent win-win from the two-way analysis) results in a 27% loss to hydropower capacity. The large number of efficient combinations, though, renders this type of tabular analysis relatively useless for comparing outcomes.

The virtues of two-dimensional simplifications are clear, as are the shortcomings. Two-way PPFs are easy to visualize and understand conceptually. They allow us to quickly understand the consequences of increasing production of one ecosystem service in terms of another. The usefulness of this exercise is questionable, though, since any effects on other services are ignored. On the other hand, higher-dimensional analysis allows us to analyze the effects to all services jointly. However, the excessive number of combinations and absence of a visual tool make meaningful

comparisons difficult. In theory, the problem could be safely reduced to two-dimensional analysis if the omitted services do not impact welfare. To learn which services have value to humans, we turn to the results of our choice experiment.

5.2 Choice experiment results: willingness to pay and latent classes

Results from the choice experiment indicate that respondents value all four of the ecosystem services. We focus first on the MNL results, which assume no variation in preferences among respondents. All coefficients are statistically significant and have the expected sign (Table 1). Generally, respondents prefer some new dam removal plan to the status quo. They like plans that deliver more Atlantic salmon, hydropower capacity, and herring, while maintaining lake shoreline. We obtain marginal values for each service in dollar terms (Figure 5) by dividing each marginal utility by the estimated cost coefficient and simulating confidence intervals (Krinsky and Robb, 1986). We present the MNL results as a baseline, but move quickly to the latent class analysis, which is both more behaviorally plausible and more informative for decision making.

The latent class results suggest a significant degree of preference heterogeneity between four groups of respondents (Table 1). We note a few differences in sign and significance between classes. First, Class 2 is the only group that derives positive utility from maintaining the status quo, while Class 3 is the only group that associates negative utility with increasing the herring population. Though seemingly strange, there is some evidence that herring may crowd out more highly valued species (FERC,

1997). In our original specification the cost parameter for Class 4 was slightly positive and statistically insignificant. In the final specification we fix the cost parameter at zero, improving model fit, which indicates a lack of attendance to the cost attribute by those respondents (Hensher et al., 2015). As a result, we are unable to calculate monetary valuation estimates for Class 4.

We can begin to paint a picture of the groups holding varied preferences for ecosystem services from dam removal by analyzing descriptive statistics (Table 2) and WTP estimates (Figure 5) by class. Class 1, for example, is more likely to be male, lower income, concerned for the environment, and regularly recreate on Penobscot rivers. Members of this class have higher WTP for hydropower and lower WTP for lake shoreline than members of any other class. Class 2 prefers maintaining the status quo. They tend to be more averse to risks and less concerned for the environment than other classes. Nonetheless, as a result of the low marginal utility they derive from income, they hold higher WTP for salmon, herring, and lake shoreline compared to other classes. Class 3 is more likely to be educated, higher income, and own or visit a ‘camp’ (a vacation cabin). They hold lower WTP for salmon and hydropower than other classes, and would be willing to pay to *reduce* the number of herring. Finally, Class 4 is more likely to be retired, highly concerned for the environment, and perceive water quality in Penobscot rivers poorly. As noted, the zero cost coefficient prevents estimating WTP, and seemingly indicates their refusal to put a monetary value on nature.

The choice experiment results contain a lot of information about values and preferences, but there are some notable shortcomings. We know, for example, that

respondents have positive WTP for services provided by both dams and dam removals. Lacking knowledge of the biophysical constraints of the system to jointly produce the services, though, we are left to speculate about a feasible and desirable balance between tradeoffs. Additionally, the latent class analysis is useful for identifying group typologies, but the WTP results are not particularly helpful for identifying areas of agreement or disagreement, and we have no way to incorporate nonmonetary values (i.e., from Class 4). Next we combine the preference data obtained from our choice experiment with the PPF.

5.3 Combining the results to identify preferred outcomes

First, we present a simple illustrative example, combining the MNL results with an arbitrarily chosen hydropower-salmon two-way PPF (Figure 6). Maximum utility is achieved by the combination of dam removals that delivers habitat to support 15,820 Atlantic salmon and generates hydropower for about 130,000 homes. Since the PPF is relatively flat around the tangency point, combinations yielding slightly more or less salmon will provide roughly the same utility. Uncertainty about the slope of the indifference curve, represented by the 95% confidence interval, indicates that anywhere along that flat portion of the PPF may actually be the most desirable. The desirability of this two-way MNL solution, though, hinges on two assumptions: first, that hydropower-salmon can accurately represent the broader tradeoffs between services from dams and services from dam removal, and second, that variations in

stakeholder preferences do not lead to variations in preferred solutions. We explore these topics in depth below.

Representing preferences for different stakeholder groups as indifference curves (Table 3) uncovers areas of agreement that were not necessarily apparent from the WTP results (Figure 5). For example, Class 3 exhibited WTP for salmon and hydropower that was about 50% lower than Class 1, yet the indifference curves for the two groups have nearly identical slopes. Similarly, Classes 1 and 2 have statistically indistinct indifference curves for lake shoreline-herring, but Class 2 exhibited WTP for both services that was at least three times higher than Class 1. Classes 1, 2, and 3 all have similar preferences for the tradeoff between lake shoreline and salmon. Finally, despite our inability to estimate WTP for Class 4, the indifference curve approach allows that group to be included in our analysis. For example, we can see that the hydropower-herring indifference curves for Class 1 and 4 are alike.

The degree to which groups agree or disagree on preferred outcomes depends not only on indifference curve slopes, but how they interact with the shape of the PPF. We can clearly see this by comparing the results for hydropower-salmon with lake shoreline-salmon (Figure 7). In terms of hydropower-salmon, apparent similarities in preferences (e.g., between Classes 1 and 3) yield different preferred outcomes when combined with the PPF. On the other hand, the lake shoreline-salmon results show that a sharp threshold in the PPF allows a set of wildly divergent indifference curves to yield near-identical preferred outcomes.

Even when two PPFs have a similar shape, indicating a high degree of correlation in their production, results can diverge if preferences are not correlated

along the same lines. Consider hydropower-herring and hydropower-salmon (Figure 8). For Class 4, the preferred level of hydropower production is 83% of maximum when paired against herring, but it drops to just 1% when analyzed against salmon. One important unifying takeaway from our two-dimensional results is that the preferred level of one ecosystem service depends on which service we choose to analyze it against.

By analyzing all four services jointly we can obtain results that do not depend on a decision about which two services we consider representative. In addition, our four-dimensional results (Table 4) allow straightforward identification of agreements and disagreements. For example, all four classes agree on the importance of maintaining lake shoreline. In addition, Classes 1, 2, and 4 are all willing to sacrifice hydropower to benefit herring and Atlantic salmon populations. The group posing the largest impediment to agreement appears to be Class 3, particularly because of their strong preference for hydropower over fish. Further, since we know Class 3 exhibits negative WTP for herring, we may be able to move toward consensus by addressing that concern in particular. The plethora of two-dimensional results, while easier to visualize and comprehend individually, provide no such clarity. Given the tradeoff between simplicity in two-way analysis versus clarity with higher dimension analysis, it ultimately becomes a question of welfare. Do we suffer a welfare loss by simplifying to two dimensions?

Our four-dimensional solution provides additional welfare gains compared to any of the two-way results (Table 5). To avoid excessive calculations, we illustrate by returning to the MNL model, which assumes no variation in stakeholder preferences.

Table 5 reiterates that each pairwise analysis yields a different outcome. In particular, when compared to the four-way solution, the two-way results tend to underestimate the preferred levels of omitted services. Despite the differences in preferred levels of each service, though, each two-way result produces a similar welfare gain. The point estimates range from \$49.23 – \$55.83 per person. Contrast this with the four-way solution, which generates a welfare gain of \$72.31 per person. Though the overlapping confidence intervals suggest the difference is statistically insignificant, it is consistent with expectations that jointly optimizing for each service improves the outcome.

6. Conclusion

In this paper we combined empirical PPFs with results from a choice experiment survey to explore heterogeneity in preferences for the ecosystem service outcomes from dam removal. The production possibilities approach to navigating ecosystem service tradeoffs is gaining traction, but requires incorporating public preferences to effectively support decision making (Cord et al., 2017). Theoretically, preferences for ecosystem service tradeoffs can be integrated using the concept of indifference curves (Cavender-Bares et al., 2015). We provide an empirical application, demonstrating how choice experiment results can be used in conjunction with PPFs to support ecosystem service policies that match the needs of society. In particular, we explore how the framework can be used to identify areas of agreement and disagreement between stakeholder groups, and the consequences of simplifying multidimensional tradeoffs to two-way analysis.

We find that jointly analyzing a choice experiment and PPF yields insights that are not apparent from individual analysis. For example, latent class analysis of our choice experiment revealed differences in WTP between groups. When represented as indifference curves and combined with the PPF, however, some of the apparent differences become points of agreement. Specifically, the magnitude of disagreements depends not only on preferences, but also on the shape of the PPF. We also note that by representing preferences as indifference curves we can include the preferences of people who refuse to value ecosystem services monetarily.

In addition, we find that care should be taken when reducing higher dimension tradeoffs to two dimensions. In PPF models for ecosystem services, correlated services are sometimes grouped or proxied onto a single axis (e.g., Mastrangelo and Laterra, 2015; Polasky et al., 2008) to ease visualization, comprehension, and computation. Comparing our pairwise two-way results, though, indicates that the socially preferred levels of services can be highly dependent on which services are omitted. Further, we find lower welfare associated with solutions derived at two-dimensions compared to four. One possibility is to derive results using all the relevant ecosystem services, but utilize two-dimensional analysis to convey findings and facilitate discussions.

Our study has a few limitations which can be addressed in future research. First, our choice experiment elicits preferences for all four ecosystem services jointly. If services were omitted from the choice experiment in a way that matched the two-way PPF analyses, respondent preferences could differ and yield potentially more consistent results (e.g., if respondents proxy the omitted services when expressing their preferences). Future research could also test the welfare consequences of

bundling services in both the PPF and choice experiment into two groups.

Additionally, we ignore potentially important spatial factors. More research is needed to develop a model that is spatially explicit in both ecosystem service production and consumption, and to understand whether welfare gains warrant the additional complexity. Nonetheless, we think combining choice experiments with PPF is a promising framework for policymakers navigating ecosystem service tradeoffs.

References

- Ager, A.A., Vogler, K.C., Day, M.A., Bailey, J.D., 2017. Economic Opportunities and Trade-Offs in Collaborative Forest Landscape Restoration. *Ecol. Econ.* 136, 226–239. <https://doi.org/10.1016/j.ecolecon.2017.01.001>
- American Rivers, 2018. American Rivers Dam Removal Database. figshare. Dataset.
- Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M., Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. *Fish. Res.*, The Scientific Basis for Management of Salmonid stocks in the British Isles 62, 143–170. [https://doi.org/10.1016/S0165-7836\(02\)00160-1](https://doi.org/10.1016/S0165-7836(02)00160-1)
- Auerbach, D., Deisenroth, D., McShane, R.R., McCluney, K.E., LeRoy Poff, N., 2014. Beyond the concrete: Accounting for ecosystem services from free-flowing rivers. *Ecosyst. Serv.* 10, 1–5. <https://doi.org/10.1016/j.ecoser.2014.07.005>
- Bennett, E.M., Peterson, G.D., Gordon, L.J., 2009. Understanding relationships among multiple ecosystem services. *Ecol. Lett.* 12, 1394–1404. <https://doi.org/10.1111/j.1461-0248.2009.01387.x>
- Cavender-Bares, J., Polasky, S., King, E., Balvanera, P., 2015. A sustainability framework for assessing trade-offs in ecosystem services. *Ecol. Soc.* 20.
- Cord, A.F., Bartkowski, B., Beckmann, M., Dittrich, A., Hermans-Neumann, K., Kaim, A., Lienhoop, N., Locher-Krause, K., Priess, J., Schröter-Schlaack, C., Schwarz, N., Seppelt, R., Strauch, M., Václavík, T., Volk, M., 2017. Towards systematic analyses of ecosystem service trade-offs and synergies: Main concepts, methods and the road ahead. *Ecosyst. Serv.*, SI:Servicing ES-EcoSummit16 28, 264–272. <https://doi.org/10.1016/j.ecoser.2017.07.012>
- Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th ed. John Wiley, Hoboken, NJ.
- Dunlap, R., Van Liere, K., Mertig, A., Jones, R.E., 2000. Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale. *J. Soc. Issues* 56.
- FERC, 2019. Complete list of active and exempt licenses.
- FERC, 1997. Final Environmental Impact Statement: License Application for Ten Existing Hydroelectric Projects in the Kennebec River Basin and License Surrender for One Existing Project. Federal Energy Regulatory Commission, Office of Hydropower Licensing, United States.

- Fox, C.A., Magilligan, F.J., Sneddon, C.S., 2016. “You kill the dam, you are killing a part of me”: Dam removal and the environmental politics of river restoration. *Geoforum* 70, 93–104. <https://doi.org/10.1016/j.geoforum.2016.02.013>
- Hanemann, W.M., 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *Am. J. Agric. Econ.* 66, 332–341. <https://doi.org/10.2307/1240800>
- Hattam, C., Böhnke-Henrichs, A., Börger, T., Burdon, D., Hadjimichael, M., Delaney, A., Atkins, J.P., Garrard, S., Austen, M.C., 2015. Integrating methods for ecosystem service assessment and valuation: Mixed methods or mixed messages? *Ecol. Econ.* 120, 126–138. <https://doi.org/10.1016/j.ecolecon.2015.10.011>
- Hensher, D.A., Rose, J.M., Greene, W.H., 2015. *Applied Choice Analysis*, Second. ed. Cambridge University Press, United Kingdom.
- Houston, R., Chadbourne, K., Lary, S., Charry, B., 2007. Geographic distribution of diadromous fish in Maine. Falmouth, Maine: U.S. Fish and Wildlife Service, Gulf of Maine Coastal Program.
- King, E., Cavender-Bares, J., Balvanera, P., Mwampamba, T.H., Polasky, S., 2015. Trade-offs in ecosystem services and varying stakeholder preferences: evaluating conflicts, obstacles, and opportunities. *Ecol. Soc.* 20.
- Krinsky, I., Robb, A.L., 1986. On Approximating the Statistical Properties of Elasticities. *Rev. Econ. Stat.* 68, 715–719. <https://doi.org/10.2307/1924536>
- Limburg, K.E., Waldman, J.R., 2009. Dramatic Declines in North Atlantic Diadromous Fishes. *BioScience* 59, 955–965. <https://doi.org/10.1525/bio.2009.59.11.7>
- Magilligan, F.J., Graber, B.E., Nislow, K.H., Chipman, J.W., Sneddon, C.S., Fox, C.A., 2016. River restoration by dam removal: Enhancing connectivity at watershed scales. *Elem. Sci. Anthr.* 4, 000108. <https://doi.org/10.12952/journal.elementa.000108>
- Martin, E., Apse, C., 2011. Northeast Aquatic Connectivity: An Assessment of Dams on Northeastern Rivers. The Nature Conservancy, Eastern Freshwater Program.
- Mastrangelo, M.E., Laterra, P., 2015. From biophysical to social-ecological trade-offs: integrating biodiversity conservation and agricultural production in the Argentine Dry Chaco. *Ecol. Soc.* 20.

- Nalle, D.J., Montgomery, C.A., Arthur, J.L., Polasky, S., Schumaker, N.H., 2004. Modeling joint production of wildlife and timber. *J. Environ. Econ. Manag.* 48, 997–1017. <https://doi.org/10.1016/j.jeem.2004.01.001>
- Nations, F. and A.O. of the U., 2001. Dams, Fish and Fisheries: Opportunities, Challenges and Conflict Resolution. Food & Agriculture Org.
- Neeson, T.M., Ferris, M.C., Diebel, M.W., Doran, P.J., O’Hanley, J.R., McIntyre, P.B., 2015. Enhancing ecosystem restoration efficiency through spatial and temporal coordination. *Proc. Natl. Acad. Sci.* 112, 6236–6241. <https://doi.org/10.1073/pnas.1423812112>
- NOAA, 2009. Biological valuation of Atlantic salmon habitat within the Gulf of Maine Distinct Population Segment Appendix A.
- Olivero, A., Anderson, M., 2008. Northeast Aquatic Habitat Classification. Boston, MA.
- Opperman, J., Royte, J., Banks, J., Rose Day, L., Apse, C., 2011. The Penobscot River, Maine, USA: a Basin-Scale Approach to Balancing Power Generation and Ecosystem Restoration. *Ecol. Soc.* 16. <https://doi.org/10.5751/ES-04117-160307>
- Polasky, S., Nelson, E., Camm, J., Csuti, B., Fackler, P., Lonsdorf, E., Montgomery, C., White, D., Arthur, J., Garber-Yonts, B., Haight, R., Kagan, J., Starfield, A., Tobalske, C., 2008. Where to put things? Spatial land management to sustain biodiversity and economic returns. *Biol. Conserv.* 141, 1505–1524. <https://doi.org/10.1016/j.biocon.2008.03.022>
- Rose, J.M., Bliemer, M.C.J., 2009. Constructing Efficient Stated Choice Experimental Designs. *Transp. Rev.* 29, 587–617. <https://doi.org/10.1080/01441640902827623>
- Roy, S.G., Uchida, E., Souza, S.P. de, Blachly, B., Fox, E., Gardner, K., Gold, A.J., Jansujwicz, J., Klein, S., McGreavy, B., Mo, W., Smith, S.M.C., Vogler, E., Wilson, K., Zydlewski, J., Hart, D., 2018. A multiscale approach to balance trade-offs among dam infrastructure, river restoration, and cost. *Proc. Natl. Acad. Sci.* 115, 12069–12074. <https://doi.org/10.1073/pnas.1807437115>
- Scarpa, R., Thiene, M., 2005. Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences. *Land Econ.* 81, 426–444. <https://doi.org/10.3368/le.81.3.426>

- Smith, S.V., Renwick, W.H., Bartley, J.D., Buddemeier, R.W., 2002. Distribution and significance of small, artificial water bodies across the United States landscape. *Sci. Total Environ.* 299, 21–36. [https://doi.org/10.1016/S0048-9697\(02\)00222-X](https://doi.org/10.1016/S0048-9697(02)00222-X)
- Stosch, K.C., Quilliam, R.S., Bunnefeld, N., Oliver, D.M., 2019. Quantifying stakeholder understanding of an ecosystem service trade-off. *Sci. Total Environ.* 651, 2524–2534. <https://doi.org/10.1016/j.scitotenv.2018.10.090>
- Train, K., 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.
- United States EIA, 2019. Maine - State Energy Profile Overview - U.S. Energy Information Administration (EIA) [WWW Document]. URL <https://www.eia.gov/state/?sid=ME> (accessed 3.26.20).
- USGS, 2017. National Hydrography Dataset (NHD) Plus, Version 2.
- Vallet, A., Locatelli, B., Levrel, H., Wunder, S., Seppelt, R., Scholes, R.J., Oszwald, J., 2018. Relationships Between Ecosystem Services: Comparing Methods for Assessing Tradeoffs and Synergies. *Ecol. Econ.* 150, 96–106. <https://doi.org/10.1016/j.ecolecon.2018.04.002>
- White, C., Halpern, B.S., Kappel, C.V., 2012. Ecosystem service tradeoff analysis reveals the value of marine spatial planning for multiple ocean uses. *Proc. Natl. Acad. Sci.* 109, 4696–4701. <https://doi.org/10.1073/pnas.1114215109>

Figure 1. The Penobscot Watershed is the largest in the state of Maine

		Status Quo No Action	New Plan A	New Plan B
ATLANTIC SALMON number that can be supported by available habitat	1,500 salmon	14,000 salmon	7,000 salmon
LIKELIHOOD that salmon will utilize newly available habitat	- 100% chance of no change to salmon	75% 25% chance of no change to salmon	50% 50% chance of no change to salmon
HYDROPOWER number of homes powered	155 thousand homes	90 thousand homes	155 thousand homes
RIVER HERRING number that can be supported by available habitat	2 million herring	5 million herring	10 million herring
LAKE SHORELINE miles lost	0 miles lost 2600 remain	300 miles lost 2300 remain	150 miles lost 2450 remain
ANNUAL COST TO YOU property tax surcharge, each year for 10 years	\$0	\$400	\$50
Which would you choose? <u>check one box ONLY</u>		No Action ☐	Plan A ☐	Plan B ☐

Figure 2. Sample choice set

Figure 3. Two-dimensional PPFs for hydropower and lake shoreline plotted against herring (inside) and salmon (outside)

Figure 4. Four-dimensional optimization results plotted in two dimensions
Note: Two-way PPF (red line) is the frontier of four-way results

Figure 5. Marginal WTP (USD)

Note: WTP cannot be computed for Class 4.

Figure 6. PPF combined with MNL results to identify utility maximizing outcome
Note: Shaded area represents 95% CI for slope of indifference curve

Figure 7. Two-dimensional hydropower-salmon and lake shoreline-salmon PPFs with indifference curves indicating preferred combinations by latent class

Figure 8. Hydro-herring PPF (inner) overlaid on hydro-salmon PPF (outer) with utility maximizing combinations by latent class

Table 1. Regression results

	MNL	Latent Class Model			
		Class 1 (39%)	Class 2 (29%)	Class 3 (24%)	Class 4 (8%)
StatQuo	-0.768*** (0.128)	-3.012*** (0.327)	1.799*** (0.573)	-4.351*** (0.992)	-2.199* (1.293)
Salmon	0.028*** (0.008)	0.012 (0.014)	0.024 (0.027)	0.021 (0.026)	0.871* (0.445)
Hydro	0.005*** (0.001)	0.005*** (0.001)	0.002 (0.003)	0.012*** (0.004)	0.005 (0.004)
Herring	0.015* (0.008)	0.023** (0.011)	0.062* (0.033)	-0.115** (0.050)	0.040 (0.061)
Lake	-0.117*** (0.023)	-0.075** (0.036)	-0.266** (0.123)	-0.508*** (0.125)	-1.460* (0.859)
Cost	-0.005*** (0.000)	-0.005*** (0.001)	-0.004** (0.001)	-0.022*** (0.004)	0 (Fixed)
LL	-1163.3	-869.1			

Notes: N= 1205. *, **, *** indicate significance at the 90%, 95% and 99% level, respectively.

Table 2. Descriptive statistics by class

	Class 1	Class 2	Class 3	Class 4
Proportion of overall sample	0.37	0.27	0.25	0.11
Male (proportion)	0.56 (0.06)	0.45 (0.07)	0.46 (0.07)	0.45 (0.11)
Income in \$1,000s (mean)	69.9 (41.9)	78.3 (62.5)	84.3 (49.3)	84.4 (64.5)
Retired (proportion)	0.26 (0.05)	0.26 (0.06)	0.23 (0.06)	0.43 (0.11)
B.A. or higher (proportion)	0.60 (0.06)	0.56 (0.07)	0.65 (0.07)	0.64 (0.1)
15+ yr resident (proportion)	0.73 (0.05)	0.70 (0.06)	0.81 (0.05)	0.82 (0.08)
Own or visit a 'camp' (proportion)	0.42 (0.06)	0.45 (0.07)	0.51 (0.07)	0.41 (0.10)
Identify as 'risk averse' (proportion)	0.24 (0.05)	0.38 (0.07)	0.27 (0.06)	0.18 (0.08)
NEP Index of environmental concern (mean)	4.53 (4.83)	2.25 (4.03)	2.96 (4.20)	5.71 (4.16)
Regular river recreational use (proportion)	0.62 (0.06)	0.50 (0.07)	0.51 (0.07)	0.45 (0.11)
Good or excellent water quality rating (proportion)	0.51 (0.06)	0.55 (0.07)	0.57 (0.07)	0.25 (0.10)
Previously aware of dam removals (proportion)	0.72 (0.05)	0.66 (0.07)	0.67 (0.07)	0.80 (0.09)

Notes: N=204. Standard deviations in parentheses.

Table 3. Two-dimensional utility maximizing combinations by class

	Latent Class Model			
	Class 1	Class 2	Class 3	Class 4
Salmon				
1,000s of fish	11.04	16.77	6.74	21.92
% of max	50%	77%	31%	100%
Hydropower				
1,000s of homes	146.87	123.36	155.06	1.45
% of max	95%	80%	100%	1%
Slope of indifference curve	-2.1	-10.0	-1.7	-168.1
95% confidence interval	(-4.1; -0.2)	(-11.4; -8.9)	(-1.8; -1.7)	(-201.5; -135.3)
Herring				
Millions of fish	12.74	15.59	0.26	13.56
% of max	78%	95%	2%	83%
Hydropower				
1,000s of homes	133.24	104.78	155.06	128.58
% of max	86%	68%	100%	83%
Slope of indifference curve	-4.3	-25.3	9.55	-7.8
95% confidence interval	(-5.0; -3.6)	(-28.8; -22.4)	(9.00; 10.13)	(-10.3; -5.3)
Salmon				
1,000s of fish	17.96	17.81	17.81	18.16
% of max	82%	81%	81%	83%
Lake shoreline				
Miles lost	3.0	1.0	1.0	11.0
% of max	<1%	<1%	<1%	2%
Slope of indifference curve	-0.16	-0.09	-0.04	-0.60
95% confidence interval	(-0.32; -0.01)	(-0.11; -0.08)	(-0.04; -0.04)	(-0.81; -0.44)
Herring				
Millions of fish	14.84	14.67	0.26	12.28
% of max	90%	89%	2%	75%
Lake shoreline				
Miles lost	30.0	25.0	0.0	2.0
% of max	6%	5%	0%	<1%
Slope of indifference curve	-0.31	-0.23	0.23	-0.03
95% confidence interval	(-0.44; -0.22)	(-0.27; -0.23)	(0.21; 0.24)	(-0.04; -0.02)

Note: Slope of indifference curve is the marginal rate of substitution between two services. Confidence intervals are simulated using Krinsky and Robb procedure with 1,000 draws.

Table 4. Four-dimensional utility maximizing combinations by class

	Latent Class Model			
	Class 1	Class 2	Class 3	Class 4
Salmon				
1,000s of fish	15.56	17.92	5.14	18.14
% of max	71%	82%	23%	83%
Hydropower				
1,000s of homes	130.03	104.78	155.06	105.58
% of max	84%	68%	100%	68%
Herring				
Millions of fish	12.20	14.67	0.26	12.74
% of max	74%	89%	2%	78%
Lake shoreline				
Miles lost	20.0	27.0	0.0	12.0
% of max	4%	5%	0%	2%

Table 5. Welfare estimates derived from MNL with two- and four- way optimizations

	Two-dimensional optimizations				Four-dimensional optimization
	Hydro Salmon	Hydro Herring	Lakes Salmon	Lakes Herring	
Salmon					
1,000s of fish	15.82	14.30	17.96	17.88	16.44
% of max	72%	65%	82%	82%	75%
Hydropower					
1,000s of homes	130.03	134.69	95.32	95.32	124.00
% of max	84%	87%	61%	61%	80%
Herring					
Millions of fish	11.92	12.34	10.90	14.19	12.87
% of max	73%	75%	66%	86%	78%
Lake shoreline					
Miles lost	106.4	89.5	3.0	18.0	20.0
% of max	21%	18%	<1%	4%	4%
Welfare gain (USD)	49.23	50.98	50.35	55.83	72.31
95% CI	(21.96; 75.90)	(25.48; 78.06)	(23.84; 78.32)	(24.05; 87.30)	(42.26; 101.29)

Notes: Shaded results represent the outcomes for the omitted services that result from maximizing utility based only on the other (non-shaded) services. Confidence intervals for welfare estimates are obtained from Krinsky and Robb procedure using 1,000 draws.