Fish, Time, and Water: Essays on Environmental Resource Trade-offs

by

Patrick Lloyd-Smith

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

Department of Resource Economics and Environmental Sociology
University of Alberta

Abstract

Understanding peoples' preferences for resources that are not bought or sold in a market is a challenging and yet often inescapable undertaking for informing decisions regarding many environmental problems. This thesis presents three studies on how people make decisions involving environmental resources. The first paper focuses on the interplay between intertemporal substitution and the value of time. I develop a structural travel cost demand model that explicitly focuses on intertemporal substitution and incorporates time constraints on behavior. The results demonstrate how getting the value of time 'right' is important for assessing welfare impacts of policies with large intertemporal substitution effects. I also find people value their leisure time heterogeneously and substantially differently from their implied wage rate and this value differs by time of year. The second paper focuses on eliciting willingness-to-accept (WTA) welfare measures in public and private good settings. Using laboratory experiments, I explore whether elicitation format, survey framing, and follow-up questions can generate more truthful responses. For public goods, I extend the mechanism perspective of incentive compatibility and provide the first empirical test of the theory in a WTA context. The findings provide support for the incentive compatibility of WTA responses for public goods as long as responses have consequences for respondents. The results of the private good experiment suggest that strategic behaviour biases are in the directions expected by theory and explicit survey framing and follow-up questions can provide useful insights. The third paper uses a stated preference survey and focuses on potential endogeneity issues with including perceived consequentiality responses in econometric models of voting behavior. The results of the study suggest that the order of the valuation and consequentiality question matters for consequentiality beliefs and that these beliefs may not be important determinants of voting behaviour, once appropriate methods to address endogeneity are used. Together, these studies contribute to the ongoing research efforts to improve the validity of nonmarket methods applied to environmental resources.

Preface

All the research reported in this thesis was approved by the University of Alberta Research Ethics Board:

- Chapter 2: “Fishery management in the Gulf of Mexico (full survey)”, No. Pro00060074, 7/10/15.
- Chapter 3: “WTA Experiments”, No. Pro00059678, 15/1/16.
- Chapter 4: “Water management in cities and towns in Alberta (provincial survey) and Canada (national survey)”, No. Pro00051054, 3/9/14.

Financial support for this thesis was provided by the Social Sciences and Humanities Research Council of Canada, Environmental Defense Fund (Chapter 2), the University of Alberta and the Smithsonian Institute (Chapter 3), and the Water Economics, Policy, and Governance Network (Chapter 4).

Chapter 2 of the thesis is coauthored with Joshua K. Abbott (School of Sustainability, Arizona State University), Wiktor Adamowicz (Department of Resource Economics & Environmental Sociology, University of Alberta), and Daniel Willard (Environmental Defense Fund).

Acknowledgements

First and foremost, I would like to thank my supervisor, Vic Adamowicz, for your guidance and mentorship of the highest quality. I could not imagine a better advisor. A special thank you to Joshua Abbott for hosting another Canuck in Arizona. I am a much better economist because of you. I appreciate the insights and thoughtful feedback from my committee members, Peter Boxall and Bruno Wichmann, and my external examining committee members, Denise Young and Catherine Kling. I would also like to thank Diane Dupont for introducing me to the world of water economics. My graduate student experience has been enriched by friends and colleagues at REES for which I am thankful. Finally, I am especially grateful to my family for the support and encouragement along this journey from afar.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the summer of 2010, the British Petroleum (BP) Deepwater Horizon oil spill released over 500 million litres of oil into the Gulf of Mexico (GOM). In addition to being the largest maritime oil spill in U.S. history, the spill also set off the most extensive natural resource damage assessment in U.S. history (Bishop et al., 2017). The damage assessment estimated that welfare losses associated with recreational uses totaled $660 million (English et al., 2015) and welfare losses associated with overall ecosystem damages totaled $17.2 billion (Bishop et al., 2017). Along with heightened awareness of the adverse impacts of human development on the natural environment, the BP spill also brought back into the spotlight the apparent necessity, and challenge, of estimating people's values for assets that are not exchanged on markets.

Understanding non-market behaviour and values has long been a focus of environmental and resource economists. A substantial amount of effort has gone into assessing the validity of both stated preference (SP) and revealed preference (RP) methods. Given the behaviour and values economists seek to study are typically not collected in existing data sources, economists often play an active hand in data collection through the use of surveys, interviews, and/or experiments. Valuation techniques, SP methods in particular, are so intrinsically linked to the data collected, that much of the criticism of the method centres on the validity of the data rather than the method or analysis itself (McFadden and Train, 2017). Simply put, the perceived unreliability of responses to contingent valuation and choice experiment questions in SP surveys has been both the most damning criticism of the method as well as the most active area of research in response. While large strides have been made in improving the validity of these methods, substantial criticisms still remain (Hausman, 2012; McFadden and Train, 2017).

RP methods such as travel cost and hedonic price models benefit from being grounded in data based on actual, observed behaviour. The perceived credibility of these methods also benefits from the lack of extensive validity tests, especially relative to SP research. For RP methods, leaving aside some important concerns about data reliability such as recall bias, larger validity concerns centre on appropriate modeling assumptions. For travel cost models, relaxing strict behavioural assumptions embedded in many discrete choice models is one means of improving the validity of these methods. Furthermore, assessing the most appropriate value of time to include in these models is another long standing question that remains open.

While there continue to be many valid critiques and shortcomings of the collective methods that constitute environmental valuation, the fact remains that in many important circumstances, they are the only appropriate means of collecting preference information, describing behaviour, and estimating welfare impacts. To paraphrase Winston Churchill, environmental valuation is the worst method of assessing peoples' preferences except for all those other methods that have been tried.

This thesis is comprised of three papers that focus on understanding how people make decisions involving environmental resources. Each paper aims to tackle important research areas in environmental valuation: how people allocate their non-market activities across time and the appropriate value of time in travel cost models, whether stated willingness-to-accept (WTA) measures can be valid, and the usefulness of consequentiality questions in assessing the validity of SP methods. The papers present results from three different primary data collection efforts involving both surveys and experiments, and in some cases combining data types and approaches.

The first paper focuses on the interplay between intertemporal substitution and the value of time. The context for this paper is the GOM recreational fisheries which have experienced a dramatic reduction in season lengths for an important target species, red snapper. Working with my co-authors, I led the design and administration of an online survey to recreational anglers in the GOM. We develop a structural travel cost demand model that explicitly focuses on intertemporal substitution and incorporates time constraints on behaviour. This novel approach to capturing intertemporal substitution addresses several limitations of the commonly used repeated discrete choice model approach and is well suited to study policy settings that requires such consideration. The use of a RP method and emphasis on modeling individual behaviour distinguishes this paper from the other two. The results of the fish paper demonstrate how getting the value of time 'right' is important for assessing welfare impacts of policies with large intertemporal substitution effects. Furthermore, we find people value their leisure time heterogeneously and substantially differently from their implied wage rate and these values differ by time of year. These findings raise concerns with the common practice of only using labor market information to value people's leisure time. The empirical application has important policy implications as U.S. federal fishery officials in the GOM are currently evaluating a pilot program that allows headboat anglers to retain red snapper outside of the traditional fishing season.

The second paper focuses on eliciting WTA welfare measures. SP practitioners have largely abandoned WTA due to perceived unreliability of questions that ask respondents for compensation amounts. I designed and conducted two laboratory experiments to investigate whether elicitation format, survey design and framing, and follow-up questions can generate more truthful responses in public and private good contexts. Both WTA experiments use time as the "good" to be valued. This good was chosen to explicitly avoid many of the issues inherent in experiments, such as house money or gift effects. For public goods, I extend the mechanism perspective of incentive compatibility formalized by Vossler et al. (2012) to the WTA context and derive analogous conditions. I provide the first empirical test of the theory in a WTA context using an experiment involving voting to accept payment to give up a public good using the single binary choice format. The findings provide support for the incentive compatibility of WTA responses for public goods as long as responses have consequences for respondents.

The private good experiment focuses on identifying and controlling for strategic behaviour. Strategic behaviour biases are found to be in the directions expected by theory and explicit framing of the questionnaire and the use of follow-up questions can provide useful insights. These findings raise potential concerns with the use of non-incentive compatible elicitation mechanisms in WTA contexts and provide an alternative approach to eliciting values in private good contexts. The results suggest that special consideration needs to be given to strategic behaviour in WTA surveys that value private goods, such as payments for ecosystem (PES) programs.

The importance of consequentiality of responses found in the public good experiment mirrors a growing body of literature that assesses these concepts in surveys. While exogenously varying the degree of consequentiality of responses is relatively easy in experimental settings, implementing these concepts in surveys faces additional challenges. The consequentiality of a survey response has typically been assessed by including a follow-up question after the valuation question that asks respondents their perceptions of the extent to which the survey results will be used by policy makers or affect decision-making.

The third paper focuses on potential endogeneity issues with including these perceived consequentiality responses in econometric models of voting behaviour.

The study uses data from a SP survey implemented to value the public benefits of reducing boil water advisories (BWAs) in the province of Alberta, Canada. A novel aspect of the study is that a split sample approach is used in the survey that varies the order of the consequentiality and valuation questions. The results of the study suggest that the order of the valuation and consequentiality question matter for consequentiality beliefs and that these beliefs may not be found to be important determinants of voting behaviour, once appropriate methods to address endogeneity are used.

The BWA study makes two contributions to the environmental valuation literature. First, the results raise issues with the common practice of including consequentiality questions after the valuation question and ignoring the potential endogeneity of these consequentiality perceptions. Second, it provides the first empirical evidence on the economic benefits of reducing BWAs in communities, which is a growing policy issue in Canada.

While the methods and applications of the papers are admittedly disparate, they share some important themes. Each paper addresses the validity of SP methods to credibly elicit preference information, albeit in different ways. Using the terminology of Bishop and Boyle (2017), validity has to do with the bias of value estimates whereas reliability has to do with the variance. In the fish paper, the SP time valuation results are compared to the commonly used income-based approaches. Important differences are found on the overall average value (level), the distribution of these values in the population (heterogeneity), and the extent to which the value of time differs by season (time-varying). While the paper presents a series of robustness checks and arguments that support the use of SP information, there is no conclusive evidence for which approach more closely matches how people actually value their time.

In the WTA paper, the emphasis is on the validity of stated WTA measures in public and private good contexts. Whereas the findings of the public good experiment support the use of consequential WTA questions, strategic behaviour by respondents identified in the private good experiment imply that eliciting WTA questions for these types of goods faces additional challenges. Perhaps more optimistically, the paper provides a contribution to the literature by demonstrating how survey framing and follow-up questions can be used to help control or at least inform potential biases from strategic behaviour.

The consequentiality paradigm that has recently taken over debates about the validity of SP methods has offered a salvation to practitioners that are forever worried about the truthfulness of responses. The BWA paper takes up the challenge of eliciting consequentiality perceptions and incorporating these beliefs into econometric models. The findings raise some sober second thoughts for the use of consequentiality questions as a panacea for SP validity.

The second theme is that all three papers highlight the importance of distinguishing between private and public goods in SP settings and using appropriate elicitation formats. Elicitation formats typically face a validity-reliability tradeoff in collecting information that is analogous to the more general bias-variance tradeoff facing many quantitative researchers. For example, the single binary choice format for public goods is a valid (i.e. unbiased) elicitation format but potentially unreliable if the sample size is not sufficient, given the larger variances associated with only collecting one piece of information per respondent. On the other hand, the choice experiment format with multiple questions collects additional preference information per individual and is typically associated with smaller variances, but these estimates may be biased. The distinction between private and public goods is most pronounced in the WTA paper where separate experiments are conducted for each type of good. The BWA paper uses data from a survey that clearly positions the valued good as public. As in the WTA public good experiment, an incentive compatible single binary choice question is posed to respondents. In the fisheries paper, the time valuation question is also a clear private good setting and an alternative elicitation format, the stochastic payment card (SPC) approach, is used. This SPC elicitation format is the same for the WTA private good experiment.

The final unifying theme of the papers is that they all shed insight on the BP spill damage assessment raised at the outset of this introduction. In addition to being conducted in the same geographic area, the fish paper demonstrates how a dynamic structural model can incorporate temporal substitution which was not considered in the BP recreation study. This model combined with the individual-specific approach used to value time provide an alternative method to assessing behavioural responses and welfare impacts of seasonal closures such as the BP spill. The BP total value study conducted to assess non-use values associated with the overall GOM ecosystem damages caused by the spill used a WTP question format in the SP survey. Conceptually, the BP spill more closely aligns with a WTA case because the spill represented a loss in environmental quality and compensation was being paid to the government (on behalf of the public). The theoretical and empirical findings of the WTA paper support the use of WTA questions in the future, although practical implementation concerns remain. Finally, the BP total value study included several steps to emphasize the consequentiality of responses but the survey did not include the typical question eliciting consequentiality beliefs.[1] The challenges raised with eliciting and modelling these consequentiality perceptions illustrated in the BWA paper suggest that this emphasis on *ensuring* consequentiality rather than *eliciting* consequentiality beliefs may be warranted.

As the world grapples with growing environmental problems and the distinction between environmental and economic issues blurs, incorporating these broader environmental concerns into economic analyses continues to be an important, policy-relevant research area. Whether it is setting appropriate Piguvian tax rates or conducting cost-benefit analyses, nonmarket valuation plays a central role in providing policy-makers with the necessary information to make decisions. The research that constitute this thesis is part of the broader, on-going efforts to improve the validity of these valuation methods.

[1]The BP total value study included a letter on official US Department of Commerce letterhead highlighting the use of the survey for policy-making and included specific text before the valuation question, "Government officials will take people's votes into account in deciding what should happen in the Gulf (U.S. Department of the Interior, 2016)

Chapter 2

Intertemporal Substitution and the Value of Leisure Time

Understanding the extent to which people substitute activities across time is important for evaluating behaviour and welfare impacts in many contexts including assessing the damages caused by oil spills and climate change impacts. We develop a structural demand model that explicitly focuses on intertemporal substitution and incorporates time constraints on behaviour. We also implement a flexible, individualized approach to measuring how people value their leisure time and how substitutable time is across periods. The model is estimated in an empirical application using data on recreation demand. The results demonstrate how getting the value of time 'right' is important for assessing welfare impacts of policies with large intertemporal substitution effects. We find people value their leisure time heterogeneously and substantially differently from their implied wage rate and these values differ by time of year. These findings raise concerns with the common practice of relying solely on labour market information to value people's leisure time.

2.1 Introduction

The ability of individuals to intertemporally substitute activities has significant implications in a number of areas of economic decision-making, including transportation choices (Davis, 2008; Arnott et al., 1993), labour supply decisions (Connolly, 2008; Shi and Skuterud, 2015), averting behaviour related to pollution (Graff Zivin and Neidell, 2014), time-use decisions (Castro et al., 2012), leisure travel (Van Nostrand et al., 2013), and recreational activities (Hartmann, 2006; Kuriyama and Hanemann, 2006). Regulations and events often impact the temporal availability and quality of these activities. Laws restrict the consumption of some goods such as alcohol to certain times of day (Boyes and Faith, 1993). Drivers respond temporally to congestion and time-of-use road charges (Hess et al., 2007). Leisure activities such as fishing and hunting are rationed using season restrictions. Oil spills and other adverse environmental events may lead to temporary closures of recreation areas. Climate change is expected to shorten the season length for winter activities such as skiing and extend the season length for some summer activities such as beach visits and golf (Mendelsohn and Neumann, 1999; Shaw and Loomis, 2008). In all these contexts, understanding how individuals substitute activities across time is critical to anticipating behavioural responses and accurately assessing costs and benefits.

Modeling how individuals allocate activities across time is closely related to how individuals value their time. Estimates for the value of time (VOT) are important for a wide range of areas in the academic literature and policy applications including transportation (Small and Verhoef, 2007), the value of a statistical life (Ashenfelter and Greenstone, 2004), monetary economics (Karni, 1973; Mulligan, 1997), understanding lifecycle consumption patterns and patterns of non-market work over the business cycle (Aguiar and Hurst, 2007; Aguiar et al., 2013), recreation demand (Phaneuf and Smith, 2005), and policy evaluation (Calfee and Winston, 1998; Bento et al., 2009). The VOT is especially pertinent for travel cost demand models because they use the costs of a trip to a site as implicit prices to evaluate behaviour and welfare. While there is a large literature illustrating how demand forecasts and welfare estimates can vary depending on the VOT used (e.g., Fezzi et al. (2014)), there is no consensus on the most appropriate approach to valuing time (Palmquist et al., 2010). By far the most common approach is to utilize some fraction of the hourly wage rate calculated from household income questions (Parsons, 2003). While this approach is grounded in the theory of the leisure-labour tradeoff, it faces a number of issues such as practical challenges in converting self-reported household income measures to hourly wage estimates, accommodating people outside the labour market, and the questionable assumption that leisure time is valued at a fixed proportion of the wage across broad cross-sections.

The assumption that the individual VOT is stable over a given temporal interval may be suspect as well. The VOT can vary depending on the time of day (Tseng and Verhoef, 2008), the amount of time available (Palmquist et al., 2010), and may vary seasonally within a year if leisure time is not perfectly fungible. The presence of household or employment constraints can limit the ability of individuals to trade leisure days at different times of years. For example, households with children may only be able to take family vacations during scheduled school breaks and employers may implicitly or explicitly limit employees' ability to take their vacation days at will. Furthermore, the structure of the traditional work schedule, with its fixed weekends and holidays, may limit fungibility as well. These predetermined leisure days typically vastly outnumber discretionary vacation days, where the individual can choose the timing.[1] Taken together, these reasons for the lack of fungibility of leisure time suggest that the VOT may depend on the time of year.

In this paper, we develop and implement a structural demand model that explicitly focuses on intertemporal substitution. We start from a static Kuhn-Tucker (KT) model which provides a utility-consistent framework for modeling decisions at both the extensive (what good to consume) and intensive (how much of a good) margins while also allowing for zero consumption levels (i.e. corner solutions) in a unified setting (Phaneuf et al., 2000; von Haefen et al., 2004; Bhat, 2008). We incorporate temporal considerations in two ways. First, we recast the choice set from 'what good to consume' to 'when to consume', allowing us to study intertemporal substitution patterns within the robust substitution framework inherent in the random utility model (RUM). Second, we relax the assumption that all leisure time can be allocated anytime, and is therefore valued equally across season, by implicitly incorporating seasonal leisure time constraints. Most demand models simply collapse the time constraint into the budget constraint under the assumption that time can be traded off against money at a constant rate - a practice that has been challenged theoretically and empirically (Shaikh and Larson, 2003; Castro et al., 2012).

We also allow flexibility in how individuals value their leisure time, rather than assuming that heterogeneity across individuals in their valuation of leisure time arises solely

[1] For example, assuming a 5 day work week and 10 days of vacation a year, the ratio of weekend days to vacation days is around 10 to 1. While vacation days provide individuals the opportunity to allocate days off when they are most valuable, the average number of paid vacation days per worker in the United States is only 10 (Ray et al., 2013), over a quarter of American workers do not receive paid time off (Bureau of Labor Statistics, 2015), and around 50 percent of available paid time is not taken on average (Glassdoor, 2014).

through variation in their wage. We use responses to money-time and leisure-season trade-off questions and estimate individual-specific VOT estimates that embed a substantial amount of both observed and unobserved constraint heterogeneity such as employment status and family life, as well as heterogeneity based on individual characteristics such as income and alternative uses of time.

We estimate the model using revealed and stated preference data collected from recreational anglers in the U.S. Gulf of Mexico (GOM). Comparing our individual-specific VOT estimates to the conventional measures, we find that the average individual-specific VOT is around 70 percent of hourly income, significantly larger than the one-third of wage rule of thumb typically implemented in the literature. More importantly, the correlation between the two estimates is small, suggesting that people value their leisure time quite differently than what their labour market returns imply. Furthermore, we find evidence of a substantial seasonal variation in the VOT. The implications for welfare impacts can be significant. Using the individual-specific approach to valuing time results in 50 to 66 percent higher welfare measures for policies with large intertemporal substitution effects but only small differences for policies with small intertemporal substitution opportunities. Thus, getting the VOT 'right' is important for assessing impacts of policies or exogenous shocks with potentially large intertemporal substitution effects.

2.2 Relevant Literature

2.2.1 Modeling Temporal Substitution

This paper contributes to the literature on how people choose to allocate activities across time.[2] Economists studying recreation demand have utilized and developed a wide range of models to address substitution across (typically spatial) choice alternatives within a temporal interval. However, modeling to incorporate substitution behaviour across time is comparatively limited (Phaneuf and Smith, 2005). Existing models that incorporate temporal considerations in recreation demand models fall into two classes: fully and partially dynamic choice models. The primary difference is that partially dynamic models are solely "backward looking" in considering the implications of past decisions for current

[2] In this paper, we directly address two temporal research needs identified by Phaneuf and Smith (2005) in their review chapter on recreation demand models: "the role of inter-temporal constraints (and opportunities) in individual choice" and "the opportunity cost of time".

welfare, whereas fully dynamic models are "forward looking" in that they allow individuals to consider the implications of their current decisions on future welfare and decision making.

Fully dynamic models include the rational habit formation model Adamowicz (1994), and dynamic programming models within a random utility maximization (RUM) framework (Provencher and Bishop, 1997; Baerenklau and Provencher, 2005). Estimation complications with fully dynamic models have led to assumptions that substantially reduce the practical usefulness of these types of models, which largely explains their lack of widespread adoption in the literature (Phaneuf and Smith, 2005; Swait et al., 2004).

The much more prevalent approach to including temporal dimensions in choice models is the use of partially dynamic recreational models as they are simpler to estimate and can easily incorporate a wide range of preference heterogeneity. These models commonly employ a repeated static RUM specification, where individuals are assumed to repeatedly make decisions of whether to take a trip to one of the recreation sites or stay home. Temporal effects can be incorporated in three different ways. First, state dependence effects across choice occasions can be incorporated through the use of variables measuring the total number of (consecutive) times a given option was chosen (Moeltner and Englin, 2004; Boxall and Englin, 2008). Second, temporal correlation and substitution patterns across choice occasions can be included through the use of repeated random parameters logit models with error components (Herriges and Phaneuf, 2002). Third, the choice set can be recast from an individual choosing different sites to choosing different time periods to take a trip as in Swait et al. (2004) and Carson et al. (2009). The main finding from these partially dynamic studies is that past experience with recreation sites matters for estimation and welfare results.

There are four main challenges associated with the use of repeated discrete choice models in a temporal context. First, the researcher needs to specify the number of choice occasions that each person faces. The usual approach is to use the same number for all individuals and set it equal to the maximum number of trips in the dataset. While Lupi (2005) has shown that the trip prediction and welfare measures are invariant to the number of choice occasions, it is not clear that these results hold if there are a heterogeneous number of choice occasions.[3] The second issue is the role of the error terms in these models. The most common specification of the repeated discrete choice models is the

[3]For example, long distance tourists may only make the decision to visit a specific beach once a year, whereas local residents may make weekly beach visit decisions. These differences in the number of choice occasions may be confounded with preference heterogeneity if the same number of choice occasions are used for each individual in the modeling.

repeated nested logit model. Each choice occasion is modelled independently, not only across individuals but also across choice occasions, which explicitly excludes the possibility of intertemporal substitution.[4] More flexible models such as random parameters logit models can capture cross choice occasion correlation patterns but place a large burden on the error terms in accounting for substitution, preference heterogeneity and all other unobserved drivers of choices. Disentangling these different roles is difficult. Third, repeated discrete choice models assume a constant marginal utility from taking trips in a period and do not incorporate satiation effects (Bockstael and McConnell, 2007). Finally, repeated discrete choice models focus on the decision process at the choice occasion of what site to choose, and no consensus has emerged on how to consistently link these individual choice occasions to decisions made over the season (von Haefen and Phaneuf, 2005).

The KT model used in this paper explicitly addresses each of these challenges. The structural econometric framework jointly models both extensive (when to take a trip) and intensive (how many trips to take) choice margins in a utility-consistent framework. No decisions on the number of choice occasions need to be made. Substitution patterns are captured through utility parameters rather than relying on unobserved error terms that also characterize unobserved preference heterogeneity. Furthermore, KT models relax the assumption of constant marginal utility of trips, utilizing choice behaviour to estimate the rate of satiation.[5]

2.2.2 The Valuation of Time

This paper also contributes to the VOT literature by providing a flexible approach to computing individual-specific VOT estimates that allows heterogeneity in how people value their time, while also allowing seasonal variability in the VOT. Estimates of the VOT are especially pertinent for travel cost demand models because, in addition to monetary costs of travelling to a site, all travel cost models require assumptions regarding how people value their leisure time. Since Cesario and Knetsch (1970) illustrated the biases to welfare estimates of recreation trips from excluding time costs, there has been

[4]For example, English et al. (2015) used a repeated nested logit model to estimate the lost recreational use from the Deepwater Horizon oil spill.

[5]Kuriyama and Hanemann (2006) modify the static KT model by extending the choice set to include multiple periods. However, their model specification does not include any time constraints and assumes that the individual cannot adjust the numeraire good across periods which limits intertemporal substitution. It is perhaps not surprising that their empirical results suggest that there are relatively small differences in the in-sample trip predictions and welfare estimates if these intertemporal considerations are ignored.

controversy in how to value leisure travel time. By far the most common practice is based on the labour-leisure trade-off and uses income information to value time. This approach is theoretically grounded in the time allocation framework of Becker (1965) and assumes that time can be transferred freely between leisure and work, implying that the monetary value in labour can be used to value leisure time. In practice, this income-based approach almost always uses a constant fraction of the hourly wage rate computed from self-reported income.[6]

While using a constant fraction of the wage rate is the most common practice, there are some complications associated with this approach. Converting self-reported household income measures to an hourly wage estimates raises a number of issues including whether to use household or personal income, how to handle non-wage income, and assumptions on the number of hours worked. Furthermore, wages may be a poor proxy for VOT for people that are outside the formal labour market such as the unemployed, the retired, or students. Implicitly the income-based measures value their time at zero. More fundamentally, the income-based approach to valuing time imposes the arbitrary assumption that all leisure time is valued at a fixed proportion of the wage.

Some alternative approaches to valuing time better reflect time's lack of fungibility. Smith et al. (1983) consider two different types of time constraints in a household production model for recreation decisions. The first is a long-run constraint where individuals divide time between labour, recreation and non-recreation activities. The second constraint is short-run in nature where individuals allocate their recreation time between alternative sites. The theoretical model illustrates the interrelationship between an individual's time constraints and the corresponding VOT, and the authors suggest collecting detailed data on the nature of individual time constraints. Feather and Shaw (1999) use a two-equation labour supply model to estimate a shadow VOT that reflects the fact that not all individuals can smoothly trade-off work time for leisure time. For flexibly employed workers, the shadow value equals the market wage, but is less than (more than) the market wage for under-employed and unemployed (over-employed). Palmquist et al. (2010) note that the shadow VOT may differ depending on the time horizon considered and propose a hierarchical decision structure. They combine the long run shadow VOT estimates from the Feather and Shaw (1999) approach with responses to short-run time/money trade-off

[6] English et al. (2015) reviewed 65 recreation demand studies that used a value for travel time, and around 50 percent used 1/3 of hourly income. The majority of the remaining studies used either zero values or the full amount of hourly income.

questions in a stated preference survey to estimate the marginal value of weekly recreational time.[7] The specific time-money trade-off question was structured as options to purchase a personal assistance service for household maintenance activities such as yard work/gardening and running errands. Palmquist et al. (2010) find that the short-run marginal value of recreational time increases as trip lengths extend from 2 to 8 hours.

Other revealed preference contexts for valuing time that do not directly rely on income are road tolls and driving behaviour. Fezzi et al. (2014) estimate the value of travel time using data on the decisions to take toll roads that save time compared to slower free roads to beaches in Italy. While noting substantial observed and unobserved heterogeneity, they estimate a mean VOT between 50 and 70 percent of household income. The advantage of this approach is that it uses actual decisions in a recreation context, but it is limited in its applicability since toll roads are not common in many parts of the world and the approach may suffer from omitted variable bias (Wolff, 2014). Wolff (2014) analyzes hourly driving speed decisions as a function of gasoline prices and finds that a one-dollar increase in the price of gas per gallon decreases speed by 0.27 miles per hour. Using this relationship, the VOT is calculated to be 50 percent of the gross wage rate.

While most applications employ a relatively simple measure of the VOT there is ample evidence that the VOT is heterogeneous and may not align with the standard 1/3 of the wage rate assumption. Furthermore, there are both theoretical arguments (DeSerpa, 1971) and empirical evidence (Tseng and Verhoef, 2008; Palmquist et al., 2010) that the VOT may differ depending on context. In this paper, we implement an approach to valuing time that incorporates a substantial amount of observed and unobserved heterogeneity while also allowing the VOT to differ by season.

2.3 Conceptual Model

We modify the traditional static KT model (Bhat, 2008; von Haefen and Phaneuf, 2005) in two ways to incorporate temporal considerations.[8] First, we redefine the choice as how many trips to take within different time periods, rather than focusing on the locations of trips as in most standard recreation demand models.[9] This re-framing of the choice set

[7]Casey et al. (1995) and Ovaskainen et al. (2012) provide two other examples of the use of stated preference questions to elicit the VOT in a recreation context.

[8]These more recent specifications of the KT model build on earlier work by Wales and Woodland (1983).

[9]This recasting of the choice set from sites to time periods is similar to the nested logit specification in Carson et al. (2009).

allows us to study intertemporal substitution patterns similarly to spatial substitution in RUMs. Second, we extend the constraint set of the KT model to include season-specific, individual-specific leisure time constraints as well as a monetary budget constraint to reflect the fact that leisure time may not be perfectly substitutable across time periods.

With these modifications, we can outline the conceptual model that underlies the empirical analysis. We take the work-leisure trade-off, and thus income, as given; assume a certain number of fixed leisure days in each (sub-annual) period; and assume the total number of vacation days can be considered exogenous variables. Each individual is assumed to maximize annual utility through choice of recreation days and non-recreation leisure days in each time period and a numeraire good over the entire year, subject to a monetary budget constraint, an annual constraint on vacation days, and a constraint on available leisure time for each time period. Individuals face the following problem[10] :

$$\max_{r_t,\ell_t,v_t,x} \sum_{t=1}^{T} U(r_t,\ell_t,Q_t,x) \quad \text{subject to} \quad \begin{aligned} & y = \sum_{t=1}^{T} c_t r_t + x, \\ & \sum_{t=1}^{T} v_t = H, \text{ and} \\ & L_t + v_t = r_t + \ell_t \;\; \forall t, \end{aligned}$$

where:
r_t is the number of recreation days at time t,
ℓ_t is the number of non-recreation leisure days,
Q_t is a vector of quality characteristics for recreation,
x is the numeraire good with price normalized to one,
y is annual income,
c_t is the monetary cost of a recreation day,
v_t is the number of discretionary vacation days,
H is the total annual number of vacation days, and
L_t is the total number of fixed leisure days such as weekends/holidays.

Appendix 2.A provides the Lagrangian for the optimization problem and resulting first-order conditions. The KT conditions that implicitly define the optimal number of recreation trips to take in time period t are given by:

[10] We ignore discounting between time periods (i.e. the discount factor equals one).

$$
\begin{aligned}
&\frac{U_{r_t}}{U_x} \leq c_t + \frac{\mu_t}{\lambda}, \quad t = 1, ...T, \\
&r_t \left[\frac{U_{r_t}}{U_x} - c_t - \frac{\mu_t}{\lambda} \right] = 0, \quad t = 1, ...T.
\end{aligned} \tag{2.1}
$$

where μ_t is the Lagrangian multiplier associated with the leisure time budget constraint in period t, and λ is the Lagrangian multiplier associated with the monetary budget constraint. The left hand side of the first equation is the Marshallian "virtual price" of r_t while the travel cost on the right hand side is composed of the monetary out of pocket costs, c_t, and the time-specific opportunity cost of a leisure day, μ_t/λ. For time periods with a positive number of recreation trips, the virtual price and travel cost are equalized, and the number of trips is chosen such that travel costs equal the marginal rate of substitution between recreation trips and the numeraire. If no recreation trips are taken in a given time period then the virtual price is bounded from above by the travel cost. The first-order conditions from the conceptual model provide estimating equations for the empirical analysis.

If fixed leisure days L_t were fungible across time periods (as are vacation days in this conceptual model) then individuals would be free to allocate this leisure time, along with vacation days, throughout the year to equalize the opportunity cost of a leisure day (i.e. $\mu_t/\lambda = \mu/\lambda$ for all t). However, there are ample reasons why this might not hold in reality. Fixed seasonal leisure days impose a positive upper bound on leisure time in periods when less leisure may actually be desired. The model predicts that no discretionary vacation time will be taken in these periods with surplus leisure and that the marginal VOT will be less in these periods than those in which vacation time is consumed.

While not explicitly included in the conceptual model, limitations on the ability to reallocate vacation time across seasons may lead to cases in which vacation time is utilized in all periods and yet the opportunity cost of leisure is not equalized across periods. For example, if an individual faces pressure at work not to take too much vacation in a month, then they may be unable to equalize the value of leisure across periods. Similarly, an individual making decisions about family leisure activities may be limited in their ability to arbitrage their allocation of vacation days due to spousal labour constraints, school schedules, etc.

2.4 Empirical Application and Data

The empirical application is an evaluation of a policy change that involves potential intertemporal reallocation of recreational headboat fishing trips in the GOM.[11] While anglers on headboat trips fish for a variety of species, two of the most important target species for marine recreational fishing in this region are red snapper and gag grouper, both of which are managed by federal fishery regulations. The recreational fishery for these species can be considered regulated open access (Homans and Wilen, 1997) with nominal state license fees, aggregate catch limits, bag and size limits, and season lengths. There are, however, no direct limits on the number of angler trips during the season. As a result of increased fishing pressure, the fishing season for red snapper in GOM federal waters has decreased substantially from around 200 days in the early 2000s to only 42 days in 2013 and 9 days in 2014. Gag grouper has followed a similar, yet somewhat muted, pattern, with fishing seasons down to five or six months in recent years.[12]

2.4.1 Survey Design and Structure

We developed an online survey to collect information on the behaviour and preferences of headboat anglers in the GOM. To evaluate the survey and ensure that questions were interpreted correctly, we conducted two focus groups with local anglers in Pensacola, Florida in August 2015. We pre-tested the online version of the survey in October 2015 with a subset of the sample to update the experimental designs and to ensure there were no technical issues.[13]

[11] A headboat is a vessel licensed to carry groups of 15 or more passengers on recreational fishing trips in the exclusive economic zone of the US GOM (Gulf of Mexico Fishery Management Council, 2016). Headboats (also called party boats), charter boats, and guide boats are different types of for-hire fishing vessels participating in GOM marine recreational fishing. For more background on the fishery, see (Abbott and Willard, 2017).

[12] Starting in 2014, the regulator, the National Oceanic and Atmospheric Administration (NOAA) Fisheries, issued an exempted fishing permit for two years to the GOM Headboat Collaborative. This pilot program used an allocation based management strategy and, instead of being constrained by short fishing seasons, the Collaborative was allocated a fixed quota of red snapper and gag grouper that could be caught anytime of the year. Thus, in 2014 and 2015 for the first time in almost 20 years, recreational anglers could fish for these target fish species year-round. A total of 19 vessels from various ports in the GOM participated in the program. The current paper is part of a broader research project that aims to evaluate the pilot program, including the benefits to recreational headboat anglers of a more flexible fishing season (Abbott and Willard, 2017).

[13] Of the 200 individuals invited by email to participate in the pre-test, 39 surveys were completed. The full survey instrument is presented in Supplementary Material B. The order of Sections 3, 4 and 5 was randomized to account for possible ordering effects.

The survey consists of six sections. The first section includes questions on the respondent's vacation behaviour, familiarity with the GOM, as well as general recreation questions. Detailed recall questions on the number and characteristics of headboat trips the respondent took in the previous year are included in the second section. A key aspect of the survey is that trip information is collected for both partial (4 to 8 hours) and full (8 to 15 hours) day trips for each of four different time periods: Winter/Spring (January to May), June, Summer (July to August), and Fall (September to December).[14]

The third section includes two contingent behaviour questions where respondents indicated the number of partial and full day trips they would have taken in the previous year under alternative fishery management regimes with varying fishing season lengths, bag limits for the target species, and prices per partial and full day headboat trip.[15] The first contingent behaviour question (Policy A) always included a "status quo" restricted fishing season (June for red snapper, July to December for gag grouper), with a bag limit of 2, but various headboat trip prices. The second question (Policy B) allowed year round fishing for the target species, a bag limit of 1, 2, or 3, and various headboat trip prices. The headboat trip price levels were 50/80, 80/130, 120/200, and 150/250 for partial/full day trips. An example of a contingent behaviour question is provided in Figure 2.B.1 in Appendix 2.B.

The fourth section includes a trip choice experiment that is not the focus of the current study. The fifth section includes the leisure time valuation questions. Respondents were presented with two choice scenarios asking them to sacrifice time for a monetary payment to either participate in a focus group or complete a short-term contract sorting paper files. Both willingness-to-accept (WTA) scenarios were for 8 hours near the respondent's home during one of their days off in the three summer months (June, July and August).[16] The choice question used a stochastic payment card (SPC) approach which combines a payment card with polychotomous choice responses. The SPC approach is closely related to the multiple-bound discrete choice (MBDC) approach but allows respondents to use a combination of words and numerical values to more easily express their preferences and uncertainty (Wang and Whittington, 2005).[17] The main advantage of the SPC approach

[14]June is given its own time period because it is the month in recent years when red snapper is allowed to be retained.

[15]The target species was either red snapper or gag grouper depending on what part of the GOM the respondent went fishing. Respondents had an 80 percent chance of receiving the gag grouper survey version (20 percent chance of red snapper version) if they took a headboat trip from Southwest Florida. All other respondents had an 80 percent chance of receiving the red snapper survey version.

[16]We used a WTA format instead of willingness-to-pay (WTP) because individuals are giving up their time to travel for recreation rather than buying their time.

[17]Other applications of the SPC approach are Bollino (2009) and Wang and He (2011).

in our context is that it efficiently gathers substantial preference information per question for deriving individual specific estimates of the VOT. An example of the focus group question is presented in Figure 2.B.2 in Appendix 2.B.

A set of leisure day trade-off questions were posed to calculate marginal rates of substitution (MRS) for leisure days throughout the year. Each respondent was presented with two questions with three alternatives each: two scenarios that changed the number of leisure days in the summer, fall, and winter periods, and a no change option.[18] Respondents chose their most and least preferred options. Figure 2.B.3 in Appendix 2.B presents an example of the leisure time trade-off question. The final section of the survey includes socio-demographic questions.

2.4.2 Survey Administration

We recruited anglers into the survey sample using respondents to an onboard survey deployed in 2014 and 2015 on headboat vessels that participated in the GOM Headboat Collaborative pilot program (see footnote 11). The onboard survey consisted of 20 questions that asked about their trip experience and collected some socio-demographic information.[19] Furthermore, respondents had an option to provide their email address if they wanted to participate in the online survey.

We administered the survey in two waves for the 2014 and 2015 samples.[20] The first wave was conducted between December 2 and 22, 2015, while the second occurred between February 11, and March 7, 2016. A total of 823 respondents completed the online survey for a response rate of 15 percent, which was the same for both waves. A total of 2,439 observations from 813 respondents are included in the KT estimation as for each individual we have the recall data and two contingent behaviour responses.[21] Table 2.1

[18]To reduce cognitive burden for respondents, we only use three time periods to calculate the marginal rates of substitution: winter/spring (January to May), summer (June to August) and fall (September to December).

[19]The 2 page survey is presented in Supplementary Material A.

[20]To provide an incentive to complete the online survey, respondents were told that everyone who completes the survey will be entered into a drawing for a $100 Amazon gift card. The survey was programmed in Qualtrics. A total of three email reminders were sent to respondents that had not completed the survey. Of the 10,719 respondents who completed the onboard survey, 5,330 unique email addresses were provided (1,574 for 2014 and 3,756 for 2015) for an initial response rate of 50 percent.

[21]A further 573 individuals started the online survey but did not finish it, for a completion rate of 59 percent. The median time to finish the survey was 32 minutes. There were 10 completed surveys that were not included in the analysis for various reasons including 6 respondents with no valid US zip code for their home address, 2 respondents that did not indicate which site they visited and hence a travel cost model could not be estimated, 1 respondent that only visited a port outside of the Collaborative program, and 1 respondent who indicated they took 290 headboat trips a year.

provides summary statistics on select socio-demographic characteristics.

Table 2.1: Socio-demographic summary statistics

Variable	Description	Mean	Min	Max
Children	Dummy variable if children in home	0.34	0	1
Home in GOM	Dummy variable if live in GOM region	0.13	0	1
Fishing experience	GOM fishing experience (years / 10)	1.42	0	7.5
Angler organization	Dummy variable if part of angler org.	0.16	0	1
Male	Dummy variable if male	0.83	0	1
College Degree	Dummy variable if hold a college degree	0.57	0	1
Age	Age of respondent	50.6	18	84
Income	Annual Income ($000s)	105	15	275
Work full- or part-time	Dummy variable if working full- or part-time	0.63	0	1
Self-employed	Dummy variable if self-employed	0.10	0	1
Bachelor degree	Dummy variable if hold a bachelor degree	0.39	0	1
Graduate degree	Dummy variable if hold a graduate degree	0.18	0	1
Household size	Number of members in household	2.80	1	9

Notes: Sample size: n = 813

2.5 Empirical Model and Analysis

2.5.1 Travel Costs

Before estimating the KT model, we first compute the costs of a fishing trip (e.g., travel costs) at the individual-level which includes the headboat trip fees as well as the costs of travelling to the GOM. For each individual, we calculate travel costs using both nonmonetary opportunity costs of time (μ_t/λ) and monetary cost information (c_t). We use two different VOT approaches to determine travel costs. The first approach ($VOT_{1/3wage}$) follows the most common practice in the literature and assumes each individual values their time at 1/3 of their hourly income regardless of time of year. The second approach allows heterogeneity and seasonality in how people value their time. This individual-specific, seasonal VOT approach (VOT_{ISS}) consists of three steps. First, using responses to the time-money WTA questions for the three summer months (June, July, and August), we estimate a random parameters logit model including both observed heterogeneity through socio-demographic characteristics and unobserved heterogeneity through random parameters (Revelt and Train, 1996).[22] Using the model parameter estimates, for each individual we obtain a distribution of VOT estimates for summer months, conditioned on the individual's observed choices and socio-demographic characteristics. The mean of the conditional distribution for each individual is then simulated to derive an individual-specific VOT estimate (Hensher et al., 2015). Second, an additional random parameters logit model is estimated using responses to the seasonal leisure trade-off questions to derive individual-specific MRS estimates of a summer day for fall and winter days. Third, these MRS values are used to rescale the summer VOT estimates to yield three VOT_{ISS} estimates per person: Winter (January to May), June/Summer (June to August), and Fall (September to December). While the $VOT_{1/3wage}$ metric varies across seasons, we also calculate a single individual-specific time-constant average VOT (VOT_{ISTC}) measure for each individual for comparison to the $VOT_{1/3wage}$ measure.[23] Appendix 2.C provides more details on the two approaches to estimating the VOT for each individual.

To capture the monetary costs of travel, we develop an expected travel cost model to account for the fact that some individuals travel long distances to go fishing in the GOM (Industrial Economics Inc, 2015; Leggett, 2015). If we followed the common practice in the travel cost literature of assuming all individuals travel by car (such that travel costs are a linear function of distance), these costs may be overestimated if individuals took a cheaper mode of travel when long distances are involved, such as flying. We estimate expected travel costs by calculating a weighted average of these driving and flying costs, where the weights are the probabilities of choosing each mode of travel. The probability of flying increases as the travel distance increases and is based on actual mode choices from the 2009 National Household Travel Survey (NHTS). The main implication of the expected travel cost model is that travel costs are a nonlinear function of distance due to the large fixed costs associated with air travel (i.e. costs per-mile decreases as the total distance traveled increases). The details of the expected travel cost calculations are provided in Appendix 2.C. Table 2.2 provides a summary of the key data sources used in the empirical analysis.

[22] Socio-demographic characteristics are incorporated as affecting the means of the random parameters and include employment status (working full or part-time, self-employed, or not working), a dummy variable whether the respondent is male, education level (some college or less, a bachelor's/associate's degree, or a graduate degree), a dummy variable if the respondent's household income is above $100,000, household size, a dummy variable for whether the respondent has children, and the age index variable.

[23] To calculate the individual-specific time-constant VOT for each individual, we use a weighted average of the seasonal VOT estimates from the individual-specific approach, where the weights are the number of months in the different time periods.

Table 2.2: Summary of data sources for model variables

Variable	Source
Number of trips (r_t)	Recall and contingent behaviour survey questions.
Monetary travel costs (c_t)	*Driving and flying costs*: Various sources, see Appendix 2.C. *Fishing trip fees*: Survey of boat operators.
Time travel costs (μ_t/λ)	*Income-based approach*: Income and hours worked survey questions. *Individual-specific approach*: Stochastic payment card question (Figure 2.B.2) and leisure day trade-off question (Figure 2.B.3).

2.5.2 Kuhn-Tucker Model

Once travel costs have been calculated, we use the first-order conditions from the conceptual model as the estimating equations in the empirical analysis. These first-order conditions along with distributional assumptions for unobserved heterogeneity provide the likelihoods for estimation. To operationalize the model, we need to define the preference specification for utility. We use the translated generalized constant elasticity of substitution (tCES) utility function (Bhat, 2008), which is closely related to the linear expenditure system (LES) used in the environmental economics literature (von Haefen and Phaneuf, 2005). This function is additively separable across fishing trips in different time periods as well as between fishing trips and the numeraire good. We assume additive separability of utility between fishing trips (r_t) and non-fishing leisure time (ℓ_t) which allows non-fishing leisure time to be included in the numeraire good term. The specific functional form is

$$U(r_t, Q_t, x) = \sum_{t=1}^{T} \frac{\gamma_t}{\alpha_t} \psi_t \left[\left(\frac{r_t}{\gamma_t} + 1 \right)^{\alpha_t} - 1 \right] + \frac{\psi_0}{\alpha_0} x^{\alpha_0} \tag{2.2}$$

where $\gamma_t \geq 0$ and $\alpha_t, \alpha_0 \leq 1$ for all t are required for this function to be consistent with the properties of a utility function (Bhat, 2008). Bhat (2008) provides a thorough overview of the interpretation of these parameters. In brief, ψ_t is the marginal utility of a trip in period t when $r_t = 0$, α_t controls the rate of diminishing marginal utility of additional recreation trips in a certain time period, and γ_t shifts the underlying indifference curves which allows for corner solutions (i.e. zero trips in a certain time period). Weak complementarity, the condition that individuals do not receive utility from a good if they do not consume it (Mler, 1974; Smith and Banzhaf, 2004), is imposed in this specification by adding and subtracting a one inside the square brackets of Equation (2.2). In our application, T=8 as we have a total of 8 fishing trip alternatives as respondents can take partial and full day trips in 4 different time periods.

The baseline marginal utility of a trip ψ_t for each trip type and time period is parameterized as an exponential function of trip quality variables Q_t to ensure $\psi_t > 0$. ψ_t also includes multiplicative omitted heterogeneity by individual and choice alternative so that $\psi_t(Q_t, \varepsilon_t) = \exp(\beta Q_t + \varepsilon_t)$. We include observed heterogeneity by interacting individual-specific variables with the quality variables, with the numeraire good serving as the base.[24]

There are a number of identification concerns in KT models that must be addressed before estimation. First, Bhat (2008) describes how γ_t and α_t both influence the quantity of good t consumed through their impact on satiation effects, such that it is difficult to disentangle these two effects. We restrict the satiation parameter to be constant across all goods ($\alpha_t = \alpha_0 = \alpha$) while allowing the translation parameter (γ_t) to vary across alternatives. Furthermore, as in other applications of the KT model (e.g. Bhat (2008)), we restrict the α parameter to be between 0 and 1 for convergence considerations. Finally, we estimate separate scale parameters for the recall and contingent behaviour data to account for any differences in the variances of the error terms.

Another unique aspect of the KT model compared to the linear-in-income discrete choice model is the role of the income constraint. In KT models, total costs on all trips and the numeraire good must add up to income. Most applications of the KT model have used only monetary income (e.g. von Haefen and Phaneuf (2005)). Larson and Shaikh (2001) illustrate the inconsistency in using a time price in recreation models with only monetary income. A practical issue with only using monetary income is that for low-income individuals, the opportunity costs of time embedded in travel costs could cause total trip costs to exceed income. We construct a total income measure for each individual that includes both monetary and leisure income to better reflect the total resources available to an individual. Leisure income is calculated using the season-specific VOT estimates for each individual and the amount of leisure time available in each time period.[25]

[24] Specifically, the ψ_1 parameter for the outside good is specified as $\psi_0 = \exp(\varepsilon_0)$.

[25] While the VOT_{ISS} approach uses the season-specific VOT estimates, the income-based approach uses the same value for all time periods. We assume that for each month, there are 64 leisure hours available (8 days x 8 hours per day). We also estimated the models with only monetary income and the welfare results are quite similar to the total income specification. The intuition for this result is that while excluding leisure income has the expected income effect on welfare estimates, a larger budget share is now spent on fishing, increasing the preference parameters for fishing trips and causing welfare estimates to increase.

To better show the connection between the conceptual model described above and the empirical model, we can substitute the total income budget constraint for the numeraire good into Equation (2.2). The total income budget constraint consists of monetary and leisure income and subtracts the expenditure spent on fishing trips, $p_t r_t$, where p_t is the full virtual price of each trip type consisting of both monetary and time costs ($p_t = c_t + \mu_t/\lambda$). Making this substitution yields

$$U(r_t, Q_t, x) = \sum_{t=1}^{T} \frac{\gamma_t}{\alpha_t} \psi_t \left[\left(\frac{r_t}{\gamma_t} + 1 \right)^{\alpha_t} - 1 \right] + \frac{\psi_0}{\alpha_0} (y + \sum_{t=1}^{T} \mu_t/\lambda(v_t + L_t) - \sum_{t=1}^{T} p_t r_t)^{\alpha_0}.$$

Using the KT conditions in Equation (2.1) and the utility function specified in Equation (2.2), Appendix 2.A derives the following estimating equations:

$$\begin{aligned}
&V_t + \varepsilon_t = V_0 + \varepsilon_0 \quad \text{if } r_t^* > 0, \text{ and} \\
&V_t + \varepsilon_t < V_0 + \varepsilon_0 \quad \text{if } r_t^* = 0, \text{ where} \\
&V_t = \beta' Q_t + (\alpha - 1) \ln \left(\frac{r_t}{\gamma_t} + 1 \right) - \ln (p_t), \text{ and} \\
&V_0 = (\alpha - 1) \ln(x).
\end{aligned}$$

To complete the econometric model structure we assume the ε error terms for an individual are distributed according to a type 1 extreme value distribution that is independent between individuals, trip types and seasons (t), and choice occasions in the survey.

There are three additional complications that require modification to the estimation approach. First, because we are using estimated VOT variables in our travel cost calculation for the individual-specific approach, we need to account for this additional source of uncertainty in the standard errors.[26] Using the individual conditional distributions from the VOT and leisure trade-off choice models, we sample 400 vectors of the estimates of the VOT across each season and individual. We then estimate the KT model using each of these vectors. The second complication is that the recall data and the responses to the two contingent behaviour questions create a total of three sets of observations per respondent. To account for the potential correlation in preferences across responses for the same individuals, we use the clustered bootstrap approach to subsume the correlation in preferences in the parameters of the model (Cameron and Miller, 2015). To implement

[26] Lew and Larson (2005) illustrate the importance of accounting for the stochastic nature of the value of time in recreation demand modeling.

this approach, we draw with replacement, a weighted (see below) sample of 813 individuals from our dataset and estimate the model using the block of all three observations per individual. This bootstrapping procedure is repeated 400 different times in combination with the 400 distinct VOT estimates. The final complication is that individuals who completed the online survey may differ from individuals in the general headboat angler population. To help address potential issues of stratification and self-selection, we use a two-stage strategy to construct survey weights.[27] These weights are used to define the probability of an individual being sampled in the bootstrap procedure.

2.5.3 Trip Prediction and Welfare Analysis

Once we estimate the KT model to recover parameters of the utility function, we use a simulation-based approach for welfare measurement and trip prediction. We define the Hicksian compensating surplus (CS^H) for a change in price and quality from baseline levels p^0 and q^0 to new levels p^1 and q^1 using expenditure functions as

$$CS^H = y - e(p^1, q^1, U^0, \theta, \varepsilon), \tag{2.3}$$

where θ is the vector of structural parameters $(\psi_t, \alpha, \gamma_t)$ and $U^0 = V(p^0, q^0, y, \theta, \varepsilon)$. Two complications arise in solving for CS^H. First, $e(.)$ depends on both interior and corner solutions for the underlying Hicksian demands in the T time period/trip length combinations and is an endogenous regime-switching function. Second, the ε's in CS^H are assumed to be unknown to the researcher, making CS^H a random variable (von Haefen and Phaneuf, 2005).

We use Monte Carlo integration techniques to simulate multiple realizations of the errors and calculate the CS^H conditional on each simulated value. There are two steps to constructing the Hicksian compensating surplus measures. In the first step we simulate unobserved heterogeneity, and the second step uses the KT model to predict how anglers

[27]The first stage aims to ensure the spatial and temporal distribution of our sample reflects the headboat angler population. We use logbook data from all headboat vessels in the GOM to calculate the percentage of anglers in each of the four seasonal periods and GOM regions (Texas, Alabama, Northwest Florida, Southwest Florida). We then compute spatial-temporal post-stratification survey weights. The second stage addresses non-response bias, where non-response includes failure to provide an email on the 2-page onboard survey or failure to complete the internet survey, by using data on characteristics from those individuals who completed the onboard survey (see Supplementary Material A), but did not complete the online survey. We weight individuals using estimated propensity scores using the following characteristics: gender, age, income, number of years fishing, how often an individual goes fishing, and where the individual lives. Weights from the two stages are multiplied together and normalized such that the sum equals the sample size.

respond following changes to prices, site closures, or other quality changes (von Haefen and Phaneuf, 2005). For the first step, we use the conditional approach to drawing error terms such that the KT model perfectly predicts the trip decisions of anglers for periods with positive trips (von Haefen and Phaneuf, 2005).[28]

Once the errors have been simulated, the structural model is used to predict behaviour under baseline and counterfactual conditions as well as the change in welfare. Deriving Hicksian demands for welfare analysis using KT models is typically complicated and the currently available methods are either enumerative (Phaneuf et al., 2000) or iterative (von Haefen et al., 2004). These procedures can be time-intensive, and we require a more efficient approach to incorporate model parameter uncertainty.[29] As a solution, we use a recently developed approach described in Lloyd-Smith (2017) that extends Pinjari and Bhat's 2011 Marshallian demand forecasting routine to simulate Hicksian demands suitable for welfare analysis. This substantially improves computational speed because it allows for closed-form welfare simulations. Lloyd-Smith (2017) provides the analytical details for the extended routine and is included as Appendix 2.F. We use the conditional approach for welfare measurement using 500 independent sets of error draws for each individual.

For trip prediction, we follow Abbott and Fenichel (2013) and use Pinjari and Bhat's (2011) demand forecast approach to simulate Marshallian demand. We use the unconditional approach for trip predictions to evaluate the in-sample fit of model specifications using the root mean squared error (RMSE) metric and to examine substitution behaviour.[30]

[28] For time periods/trip types where trips are zero, the errors are simulated from a type I extreme value distribution that has been truncated to reflect this choice. Thus the conditional approach uses observed behaviour by individuals to characterize unobserved heterogeneity.

[29] To incorporate parameter uncertainty, we repeat the error simulations and demand predictions for each realization of the model parameters generated by each of the 400 bootstraps. Thus, for the 400 bootstrap iterations over 813 observations and 500 error realizations, we need 1.63×10^8 demand simulations for each policy.

[30] The unconditional approach uses unconditional draws from the entire distribution of unobserved heterogeneity. We use the unconditional approach for trip prediction as it does not make sense to use the conditional approach because errors are drawn such that trips are perfectly predicted. The root mean squared error is calculated over the recall data only and uses the actual trips taken by anglers and the predicted Marshallian demands from the KT model.

2.6 Results

We present three sets of results. First, we compare the individual-specific and income-based VOT estimates and show the implications for overall travel costs. Second, we present the KT model results using the two VOT approaches and the trip prediction metrics. Lastly, we present the behavioural substitution and welfare implications for three different policy scenarios.

2.6.1 Value of Time

We first compare the individual-specific VOT estimates using responses from the time valuation questions to the more traditionally used income based VOT measures.[31] Using the individual-specific approach, we calculate VOT_{ISTC} to be \$27 (median \$23) with a range of \$6 to \$96. Using the income-based approach, the average hourly equivalent wage of respondents is \$39 (median \$34) with a range of \$0 to \$429. Comparing averages, the VOT_{ISTC} measure is around 70 percent of the average hourly wage, which is comparable to the results in Palmquist et al. (2010). The convention in the recreation demand literature is to use 1/3 of computed hourly wages, which yields an average $VOT_{1/3wage}$ estimate of \$13. Figure 2.1 shows a scatterplot of the relationship between the VOT_{ISTC} and $VOT_{1/3wage}$ estimates for each individual along a 45-degree line. There is only a weak positive relationship between the two estimates (ρ=0.14).[32] Furthermore, a large number of unemployed or retired individuals are imputed a \$0 VOT using the income-based approach, yet their choice behaviour reveals a positive valuation.[33]

2.6.1.1 Marginal Rate of Substitution of Leisure

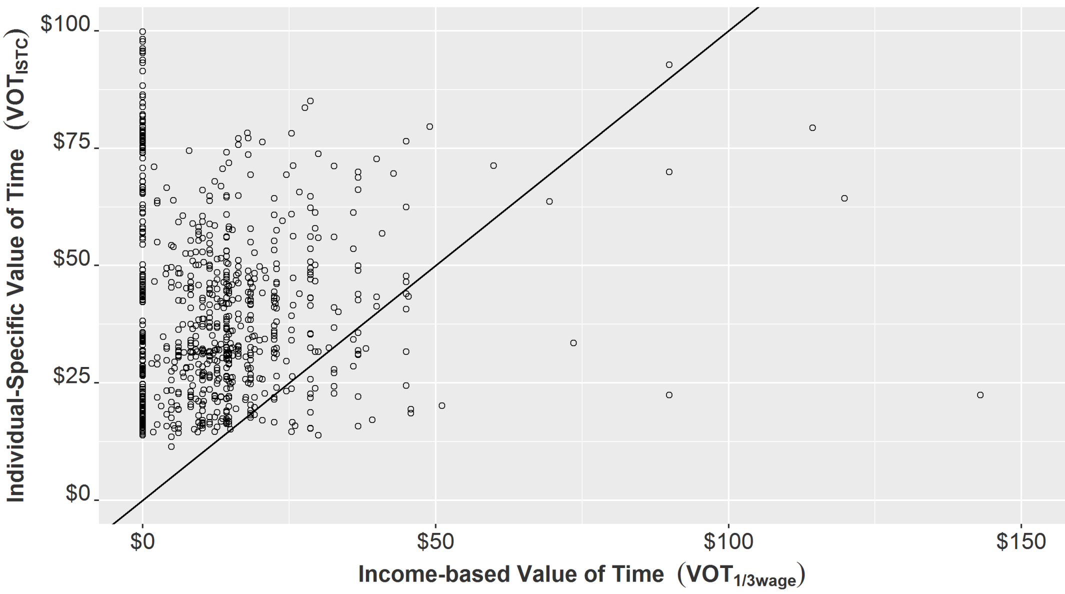
Next, we present the results on the marginal rate of substitution of fall ($MRS_{f,s}$) and winter ($MRS_{w,s}$) days for a summer day which are used to calculate the seasonal variation in the VOT. The individual-specific estimates for $MRS_{f,s}$ range from 0.29 to 1.59 with a mean of 0.65 whereas the $MRS_{w,s}$ ranges from 0.23 to 1.67 with a mean of 0.64. Figure 2.2 presents the distribution of individual-specific estimates of MRS for all respondents. If all

[31]The individual-specific estimates are calculated as the mean of the conditional distribution of the VOT using observed choices and socio-demographic characteristics. The detailed derivation of these measures at the individual level is explained in Appendix 2.C.

[32]If we only include employed individuals, the correlation coefficient is 0.22.

[33]Appendix 2.D presents a set of robustness checks on our individual-specific VOT approach; results from these checks support the validity of the approach used in this paper.

Figure 2.1: Relationship between value of time estimates per hour using individual-specific and income-based approaches



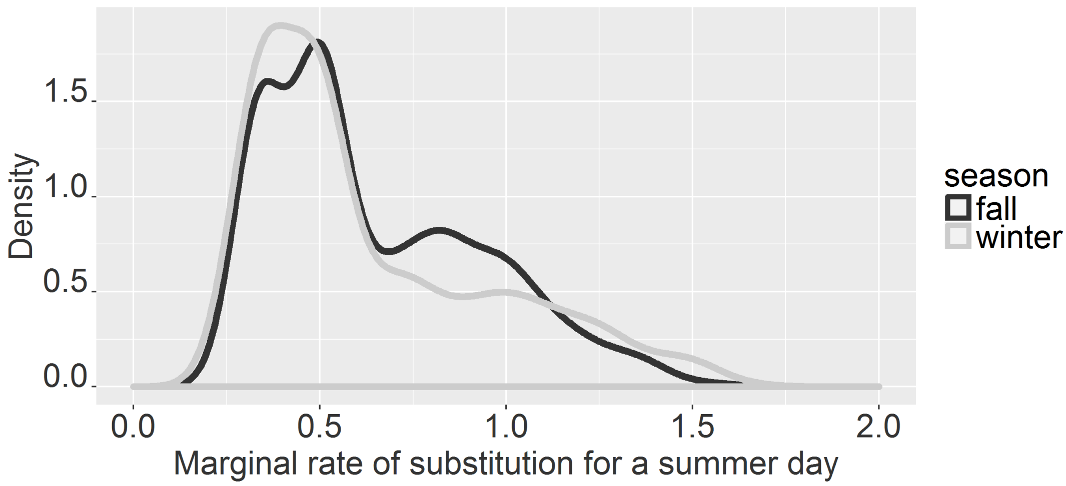
Notes: This figure plots the value of time measures from the individual-specific and income-based approaches to valuing time. Each dot represents a single individual. The 45 degree line indicates a perfect correspondence in estimates between the two approaches.

leisure time throughout the year were valued the same, then the MRS would equal one. Values less than one imply that people value summer days more than fall/winter days while values greater than one suggest the opposite. These results suggest that people do in fact value leisure time differently depending on the season, although at the population level the opportunity cost of time in the fall versus winter relative to the summer is quite similar, and there is a substantial degree of heterogeneity.

2.6.1.2 Travel Costs

The preceding results show that there are differences in the VOT estimates between the two approaches. To summarise the effects of the two VOT estimates on travel costs, we compare the time portion of travel costs to the monetary costs. The average time costs per trip using the income-based approach are \$390, or \$130 if 1/3 of this amount is used. This compares to an average of \$274 per trip using the VOT_{ISS} approach. Total monetary costs are composed of headboat trip fees, which average \$145, and costs associated with traveling to the port such as gas, accommodation, and airfare, which average \$135 per trip. The correlation of overall travel cost (including time costs) utilizing the two alternative approaches to time valuation is 0.69. Consequently, we conclude that the time portion

Figure 2.2: Distribution of individual-specific estimates of marginal rates of substitution (MRS) of fall and winter for summer leisure days



Notes: This figure plots the distribution of individual estimates of the marginal rates of substitution of fall (black line) and winter days (grey line) for a summer day. Values less than one indicate summer leisure days are preferred to fall/winter days and values more than one indicate that fall/winter leisure days are preferred to summer days.

of travel costs is a significant component of overall travel costs and differences between the income-based and individual-specific VOT approaches may have consequences for modeling.

2.6.2 Estimation

Table 2.3 reports parameter estimates for the KT model using the two alternative travel cost measures. Model 1 uses the individual-specific, seasonal (VOT_{ISS}) approach while Model 2 uses the conventional 1/3 wage income-based ($VOT_{1/3wage}$) approach. The estimated average log-likelihood at convergence of Model 2 is less than that of Model 1 suggesting the VOT_{ISS} specification fits the data better. The in-sample trip prediction metrics also report a slight improvement using the VOT_{ISS} specification. Nevertheless, both likelihood criteria and in-sample fit suggest that the improvements from utilizing the VOT_{ISS} approach are relatively modest.

Table 2.3: Parameter estimates for Kuhn Tucker model

	Model 1: Individual-specific, seasonal value of time (VOT_{ISS})		Model 2: Income-based value of time ($VOT_{1/3wage}$)	
	Estimate	z-stat	Estimate	z-stat
Marginal utility of trip parameters (ψ_t)				
Constant	-7.05	-25.92	-6.89	-24.52
Full day trip	0.00	-0.02	0.01	0.06
Winter	-0.48	-2.31	-0.41	-1.91
June	1.36	5.81	1.14	4.78
Summer	1.25	4.67	0.99	3.48
Contingent behaviour	0.48	9.85	0.52	10.63
Retain fish	0.16	2.54	0.17	2.52
Retain fish*Red snapper	0.00	0.00	-0.01	-0.16
Children	-0.34	-3.28	-0.42	-3.92
Children*Summer	0.28	2.42	0.30	2.64
Fishing experience	0.04	1.26	0.06	1.67
Angler organization	0.22	1.68	0.24	1.64
Male	0.19	1.96	0.14	1.42
College Degree	-0.36	-4.47	-0.43	-5.00
Age index	-0.12	-0.47	-0.26	-1.07
Age index*Full day trip	0.16	0.72	0.19	0.81
Age index*Winter	0.65	3.09	0.62	2.84
Age index*June	-1.10	-4.65	-1.10	-4.79
Age index*Summer	-1.05	-4.18	-1.05	-4.14
Home in GOM	-1.07	-7.86	-0.89	-6.44
Home in GOM*Full day trip	0.63	4.96	0.64	4.92
Home in GOM*Winter	-0.03	-0.25	-0.06	-0.50
Home in GOM*June	-0.14	-1.12	-0.08	-0.68
Home in GOM*Summer	-0.14	-1.02	-0.04	-0.31
Satiation Parameter (α)	0.00[a]	1.01	0.00[b]	1.05
Translation Parameters (γ_t)				
Partial day trips				
Winter (γ_1)	1.41	26.92	1.40	27.83

Continued on next page

Table 2.3 – *Continued from previous page*

	Model 1: Individual-specific, seasonal value of time (VOT_{ISS})		Model 2: Income-based value of time ($VOT_{1/3wage}$)	
	Estimate	z-stat	Estimate	z-stat
June (γ_2)	1.22	28.15	1.22	27.77
Summer (γ_3)	1.16	31.34	1.15	30.98
Fall (γ_4)	1.47	22.63	1.46	22.20
Full day trips				
Winter (γ_5)	1.69	20.22	1.69	20.00
June (γ_6)	1.26	25.72	1.27	23.19
Summer (γ_7)	1.31	24.30	1.31	22.96
Fall (γ_8)	1.64	21.47	1.64	20.80
Scale Parameters				
Contingent behaviour scale	0.85	44.84	0.86	42.42
Recall scale	0.91	47.59	0.90	46.09
N	2,439		2,439	
Log-likelihood	-18,713		-18,738	
In-sample RMSE	8.28		9.26	

Notes: z-stats are calculated using cluster bootstrap standard errors. The bootstrapping procedure is repeated 400 different times. GOM = Gulf of Mexico. RMSE = root mean squared error. Age index is calculated as the age of the respondent divided by the mean age. College Degree is a dummy variable if a respondent holds a bachelor or graduate degree. [a] The value of 0.00 was the estimated value at convergence. [b] The value of 0.00 was the estimated value at convergence.

The model coefficients are quite similar between the two VOT specifications except for the June and Summer time period dummy variables, which are higher in Model 1. Holding other quality variables and travel costs constant, full day trips are preferred to partial day trips for people who have a home in the GOM but not for people visiting from further away. Trips in June and Summer time periods are more preferred to other times of year in general, but older individuals prefer the Winter and Fall periods. The interpretation of the positive and significant coefficient for the contingent behaviour dummy variable is not straightforward because the larger estimated scale parameter for the recall responses dampens this effect.[34] The retain fish variable captures whether red snapper or gag

[34] A common finding in recreation demand modeling is to find more trips taken in the contingent behaviour scenario compared to the recall scenario (Englin and Cameron, 1996). In our data, the

grouper could be kept by the angler once caught in the certain time period and has the expected positive sign.[35] For socio-demographic variables, individuals with children are less likely to take headboat trips in general but are more likely in the summer months when school is out of session. The translation parameters (γ_t) influence the rate of satiation and the propensity toward corner (zero) solutions for a given trip type and season. In general, the greater the value of γ_t the less an individual satiates on that choice and the less likely they will choose zero trips. However, these parameters are difficult to directly compare because the time periods have a different number of months. Thus, while the larger values of the translation parameters for Winter and Fall time periods seemingly reflect lower satiation for these time periods, on a per-month basis, June actually has the lowest rate of satiation not surprisingly given that this is the month where the red snapper season is typically open. The scale parameter estimates suggest that the variance in the recall responses is greater than the variance in the contingent behaviour data.

2.6.3 Substitution and Welfare Analysis

To simulate behavioural responses and welfare impacts, we use the actual trip data for each individual as the baseline, where the average annual number of trips per angler is 3.6 and mean total trip expenditures are $1,024. We compare three policy scenarios that differ in the degree to which they bear upon intertemporal substitution possibilities:

- Policy 1: an increase in per trip prices of $25/$50 for partial/full day trips;
- Policy 2: closure of all fishing in the Summer time period; and
- Policy 3: closure of all fishing in the Fall time period.

Because Policy 1 consists of relatively small price increases (4 to 9 percent) across all time periods, incentives for intertemporal substitution should be limited. Policies 2 and 3 concern possible hypothetical temporary closures of the GOM recreational fishery caused by events such as an oil spill or the regulatory environment in response to overharvesting.

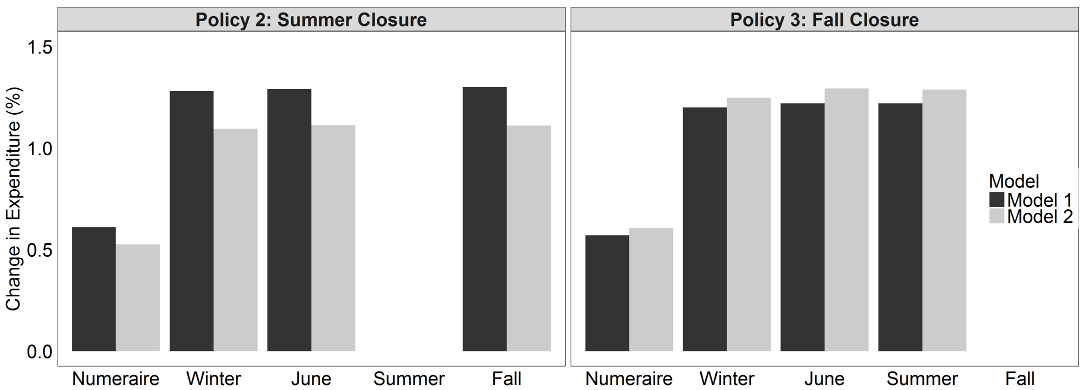
average annual number of trips per individual is 3.56 in the recall scenario and 3.65 in the contingent behaviour scenarios. As a robustness check, we also estimate both models using only the recall data. The parameter estimates are similar to the models presented in the paper using both recall and contingent behaviour data and the welfare impacts are slightly higher. This result gives us more confidence that the contingent behaviour dummy variable and heterogeneous scale parameters are appropriately controlling for any systematic differences between the recall and contingent behaviour data.

[35]For the recall responses, we used the respondent's knowledge of the GOM Headboat Collaborative pilot program to code the retain variable as either year round if they knew about the program, or only for time periods that align with the traditional target species (either red snapper or gag grouper) season if they did not know about the program.

As an example, on April 20, 2010, an explosion on the drill rig Deepwater Horizon led to the closure of much of the GOM fishery for that summer.[36] The seasonal closures of Policies 2 and 3 should generate more substitution of trips across time. We chose these specific policies to compare contexts with varying degrees of intertemporal substitution.

Before discussing welfare impacts, we first present the behavioural substitution responses to the two seasonal closure policies. Figure 2.3 illustrates the mean percentage change in expenditures on fishing trips in the different periods as well as the numeraire for the two models. Starting with the intensive margin changes between trips in different time periods, Model 1 predicts larger percent changes in expenditures than Model 2 for summer closures and the models predict quite similar percent changes in expenditures for fall closures. The KT specification, while still somewhat constrained by its additively separable structure, relaxes the independence of irrelevant alternatives (IIA) assumption embedded in type I extreme-value discrete choice models within the model parameters, rather than through the error structure (Vasquez Lavin and Hanemann, 2008). Indeed, the substitution patterns suggest small deviations from proportionate shifts in expenditures across seasons. Examining the extensive margin change, the percentage increase in numeraire expenditure is relatively modest, because fishing trip expenses are small part of total expenditures for the majority of individuals. However, the percentage decrease in total trip expenditures is estimated to be -21 percent (Model 2) to -23 percent (Model 1) under Policy 2 and -23 percent (Model 1) to -24 percent (Model 2) under Policy 3.

Figure 2.3: Behavioural responses to seasonal closure policies



Notes: This figure illustrates the average change in seasonal trip and numeraire expenditures as a result of a summer fishing closure (Policy 2) and a fall fishing closure (Policy 3) for Model 1 and Model 2.

Table 2.4 reports annual Hicksian welfare estimates per anglers for the three policy scenarios using three alternative methods for calculating the VOT. We first discuss Model

[36]The damage assessment for the Deepwater Horizon Oil Spill estimated that the welfare losses associated with recreational uses totalled \$660 million (English et al., 2015).

1 and Model 2 and then introduce Model 3. For Policy 1, the welfare impacts are quite similar between the two models, with slightly higher estimates using Model 1 (-$124) compared to Model 2 (-$119). These initial results suggest that for policies with small intertemporal substitution possibilities, the specific VOT estimate may not matter much. However, there is a large divergence between the model results for fishery closures, Policy 2 and 3. The summer fishing closure in Policy 2 results in a 66 percent larger welfare impact (-$193 versus -$116) using the VOT_{ISS} estimates compared to the $VOT_{1/3wage}$ approach. This difference in welfare estimates for the fall closure scenario is 50 percent (-$216 versus -$144), with the VOT_{ISS} estimates again yielding larger impacts. Thus, for policies with large intertemporal substitution effects, the two VOT approaches lead to sizeable differences in welfare impacts, with the conventional income-based approach underestimating these impacts.[37]

Table 2.4: Welfare estimates

Policy Scenario	Model 1: Individual-specific, seasonal value of time (VOT_{ISS})	Model 2: Income-based value of time ($VOT_{1/3wage}$)	Model 3: Individual-specific, time-constant value of time (VOT_{ISTC})
	Mean welfare impacts ($/person/year)		
Policy 1: Trip fee increase all year ($25 partial/$50 full day)	-$124 (0.21)	-$119 (0.28)	-$124 (0.20)
Policy 2: Summer fishing closure	-$193 (4.4)	-$116 (2.7)	-$162 (3.4)
Policy 3: Fall fishing closure	-$216 (5.1)	-$144 (3.5)	-$228 (5.0)

This table reports mean annual welfare impacts per individual of the three policies. Estimates are generated with 500 conditional error draws per individual. Cluster bootstrap standard errors in parentheses.

The seasonal variation in the VOT also has implications for the welfare impacts of closures in different times of year. For Model 2 and Model 3, which both use a time-constant VOT, closures in the fall have a 24 and 41 percent larger welfare impacts compared to closures

[37]The welfare estimates are derived using the different parameter estimates from the two models as well as using the different VOT measures in the welfare calculations. We can decompose the role these two drivers have on the welfare estimates by simulating welfare measures using Model 1's parameter estimates and the $VOT_{1/3wage}$ measures and vice versa with Model 2's parameter estimates and VOT_{ISS} measures. Appendix 2.E illustrates that the differences in welfare estimates are largely driven by differences in the VOT measures used to simulate welfare rather than differences in the estimated preference parameters across the two models.

in the summer. Conversely, for Model 1, which allows for seasonality in the VOT, the fall fishing closures have a 12 percent larger welfare impact compared to the closure in the summer. These results suggest the timing of the fishing closure matters for welfare.

However, there are two key differences between the VOT measures for Model 1 and Model 2: flexibility in how people value their time and seasonal variation. To isolate the impacts for welfare estimates of allowing seasonality in the VOT we also estimate the KT model and simulate welfare impacts using the individual-specific, but time constant VOT that does not include any seasonal variation in the VOT (VOT_{ISTC}). Model 3 provides the welfare results for this model. The estimated mean welfare impact is -\$162 for a summer closure compared to -\$228 for a fall closure. Comparing these impacts to Models 1 and 2, it is clear that incorporating flexibility in how people value their time is critical for closing the gap between these two models, and is a larger factor in welfare impacts compared to temporal heterogeneity. Nevertheless, incorporating seasonality in the VOT has important implications for determining how the timing of the closure matters for welfare.

2.7 Conclusion

In this paper, we develop a structural demand model that places intertemporal substitution at its core. We also compare two alternative approaches to valuing time: a flexible approach that incorporates individual heterogeneity and seasonality, and the conventional income-based approach. We implement the model using revealed and stated preference data from an online survey of recreational anglers in the GOM. We find that the individual-specific VOT is around 70 percent of hourly income, which is larger than the value of 1/3 of hourly income that has dominated applications. More importantly, the correlation between the two estimates is small, suggesting that people value their time quite differently than what their labour market returns imply implying that an income-based approach may suffer from serious measurement error. Furthermore, we find evidence of a substantial seasonal variation in the VOT. For welfare impacts, using the more flexible approach to valuing time results in 50 to 66 percent higher welfare measures for policies with large intertemporal substitution effects but only small differences for policies with small intertemporal substitution opportunities. Thus, getting the VOT 'right' is important for assessing impacts of events or policies with large potential intertemporal substitution effects such as oil spills, climate change impacts, or firm or government policies to temporally ration goods or services.

The research has two broader implications beyond modeling recreation demand. Our paper illustrates an approach to structurally modeling intertemporal substitution that is significantly more tractable than alternatives such as dynamic discrete choice models. The most common approach of using repeated discrete choice models in contexts with significant intertemporal substitution may not be appropriate because these models ignore or largely downplay the role of intertemporal substitution in decision making. The KT model allows substitution patterns to be captured explicitly through utility parameters, incorporates satiation effects in each period, and jointly models consumption decisions along both the extensive and intensive margins. Our modeling also provides an illustration of how survey information can be used within the KT model structure to augment revealed preference data to better identify key parameters.

We contribute to the VOT literature by not only undermining the "1/3 wage" rule of thumb but by also revealing that the majority of individual-level differences in the VOT are not explained by reference to their return on the labour market. Even people who are disconnected from the labour market, such as the unemployed or retired people, do positively value their time. While our individual-specific approach implicitly assumes respondents are answering questions truthfully, the income-based approach also typically relies upon self-reported data. Thus the choice between approaches cannot rest on a preference for revealed preference data alone. As detailed in Appendix 2.C, the income-based approach also requires a substantial number of assumptions to transform annual household income into an hourly wage metric. Furthermore, our more flexible approach does not preclude people from valuing their time using labour market returns. Another reason to move beyond the income-based approach is that it does not allow seasonal variation in the VOT. The results of this paper suggest that the VOT can differ depending on the season, and accounting for this variation is important for assessing welfare impacts of policies that cause intertemporal substitution of demand. Demand modeling is typically done at the individual level and thus requires valid information at this level. The flexible, individualized VOT approach used in this paper provides a useful means of obtaining this type of information. Adding a small number of questions to a survey to elicit this information seems like a small price to pay. A valuable research agenda would be to investigate the heterogeneity in the VOT and its seasonal variation to better understand what individual characteristics drive these results.

Appendix 2.A: Additional details on conceptual and empirical model

Lagrangian function and first-order conditions for conceptual model

Before constructing the Lagrangian function, we can simplify the problem by substituting the vacation constraint into one of the leisure time period constraints and removing the vacation day choice variable so that individuals are choosing only recreation and non-recreation days in each period. The Lagrangian equation is then given by

$$L = \sum_{t=1}^{T} U(r_t, \ell_t, Q_t, x) + \lambda \left[y - \sum_t c_t r_t - x \right] + \sum_t \mu_t \left[L_t + v_t - r_t - \ell_t \right]$$

We assume that the numeraire good and non-recreation leisure days have positive demand and thus the constraints are always binding and the associated Lagrangian multipliers are positive. The resulting Kuhn-Tucker first-order conditions are

$$\begin{aligned}
\frac{\partial L}{\partial r_t} &= U_{r_t} - \lambda c_t - \mu_t \leq 0, \;\; r_t \geq 0, \;\; r_t \frac{\partial L}{\partial r_t} = 0, \;\; t = 1, ...T, \\
\frac{\partial L}{\partial x} &= U_x - \lambda = 0, \\
\frac{\partial L}{\partial \ell_t} &= U_{\ell_t} - \mu_t = 0, \;\; t = 1, ...T, \\
\frac{\partial L}{\partial \lambda} &= y - \sum_t p_t r_t - x = 0, \\
\frac{\partial L}{\partial \mu_t} &= L_t + v_t - r_t - \ell_t = 0, \;\; t = 1, ...T - 1, \\
\frac{\partial L}{\partial \mu_T} &= L_T + H - \sum_t^{T-1} v_t - r_T - \ell_T = 0.
\end{aligned}$$

From the second first-order condition we know $U_x = \lambda$. We can divide the first first-order condition by λ to yield:

$$\begin{aligned}
\frac{U_{r_t}}{\lambda} &\leq c_t + \frac{\mu_t}{\lambda}, \;\; t = 1, ...T, \\
r_t \left[\frac{U_{r_t}}{\lambda} - c_t - \frac{\mu_t}{\lambda} \right] &= 0, \;\; t = 1, ...T.
\end{aligned}$$

Substituting U_x for λ in the denominator when U_{r_t} is the numerator results in Equation (2).

Derivation of estimating equations

Using the functional form for the ψ parameters described in the text and the α identifying restriction ($\alpha_t = \alpha_1 = \alpha$), the utility function specification in Equation (2) can be written as

$$U(r_t, Q_t, x) = \sum_{t=1}^{T} \frac{\gamma_t}{\alpha} \exp(\beta' Q_t + \varepsilon_t) \left[\left(\frac{r_t}{\gamma_t} + 1 \right)^{\alpha} - 1 \right] + \frac{1}{\alpha} \exp(\varepsilon_1) x^{\alpha}.$$

The partial derivative of the utility function with respect to a recreation trip and the numeraire good is equal to

$$\begin{aligned} U_{r_t} &= \frac{\gamma_t}{\alpha} \exp(\beta' Q_t + \varepsilon_t) \frac{\alpha}{\gamma_t} \left(\frac{r_t}{\gamma_t} + 1 \right)^{(\alpha-1)} = \exp(\beta' Q_t + \varepsilon_t) \left(\frac{r_t}{\gamma_t} + 1 \right)^{(\alpha-1)}, \quad \text{and} \\ U_x &= \lambda = \frac{1}{\alpha} \exp(\varepsilon_1) \alpha x^{(\alpha-1)} = \exp(\varepsilon_1) x^{(\alpha-1)}. \end{aligned} \tag{2.A-1}$$

We can manipulate the KT conditions in Equation (1) to yield:

$$U_{r_t} \leq \left(c_t + \frac{\mu_t}{\lambda} \right) \lambda, \quad t = 1, ...T.$$

We substitute in the expressions for U_{r_t} and λ from Equation (2.A-1) and use the full virtual price term p_t, where $p_t = c_t + \mu_t/\lambda$, to yield

$$\exp(\beta' Q_t + \varepsilon_t) \left(\frac{r_t}{\gamma_t} + 1 \right)^{(\alpha-1)} \leq (p_t) \exp(\varepsilon_1) x^{(\alpha-1)}, \quad t = 1, ...T.$$

Taking logarithms of both sides yield the estimating equations as

$$\begin{aligned} &V_t + \varepsilon_t = V_1 + \varepsilon_1 \quad \text{if } r_t^* > 0 \\ &V_t + \varepsilon_t < V_1 + \varepsilon_1 \quad \text{if } r_t^* = 0, \quad \text{where} \\ &V_t = \beta' Q_t + (\alpha - 1) \ln \left(\frac{r_t}{\gamma_t} + 1 \right) - \ln(p_t), \quad \text{and} \\ &V_1 = (\alpha - 1) \ln(x). \end{aligned}$$

Bhat (2008) details how these equations are used in estimating the KT model.

Appendix 2.B: Online survey sample questions

Figure 2.B.1: Sample contingent behaviour question

In recent years recreational anglers could only retain red snapper during a 1 to 1.5 month season starting June 1st. The season length and bag limit of a typical red snapper season in the recent past are presented in Policy A below.

Policy A

Season when red snapper can be retained	June
Red snapper bag limit	2
Price of one partial day (4-8 hrs) headboat trip	$80
Price of one full day (8-15 hrs) headboat trip	$130

If the Gulf of Mexico red snapper fishing policies were as described in Policy A, how many headboat trips would you have taken in 2015 in the different seasons? In considering your responses, please assume that any features about the fishing trips that are not mentioned such as sea conditions, the quality and size of the boat, the number of passengers, and bag limits and regulations for other species are the same as your 2015 experience.

2015 Gulf of Mexico Headboat Trips under Policy A

	January to end of May Holidays: New Year's, Spring break, Easter, Memorial day	June	July to end of August Holidays: Independence Day	September to end of December Holidays: Labor Day, Columbus Day, Thanksgiving, Christmas
Number of **partial day (4-8 hrs)** headboat trips in 2015 under Policy A				
Number of **full day (8-15 hrs)** headboat trips in 2015 under Policy A				

Figure 2.B.2: Sample willingness-to-accept time valuation question

Many research companies pay people to participate in a focus group.

Suppose you've been given the opportunity **to be paid a certain amount of money to participate in a full day (8 hours) focus group** near your home during one of your days off during the three summer months (June, July and August).

How likely is it you would participate in the focus group if the payment amount is…? Please select a response for each payment amount.

	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
$50					
$100					
$200					
$400					
$700					

Figure 2.B.3: Sample leisure time trade-off question

We are now going to ask you to choose between two options that change the number of leisure days between different seasons of the year.

Suppose you were given two options relative to your current situation in terms of the timing of your leisure days. Which of the options below do you *most* and *least* prefer?

Time period	**Option A**	**Option B**	**Option C**
January to May	No change	1 more leisure day	No change from your current situation
June to August	1 less leisure day	1 less leisure day	
September to December	1 more leisure day	No change	

	Option A	Option B	Option C
I **most** prefer...			
I **least** prefer...			

Appendix 2.C: Detailed travel cost calculations

In this appendix, we describe the approach to calculating the travel costs of a headboat fishing trip in the GOM. The cost of traveling to each port is calculated using both non-monetary opportunity costs of time represented as μ_t/λ in the conceptual model and monetary cost information represented as c_t. We first describe the procedure to estimate the VOT using the income-based and individual-specific approaches and then outline the approach to calculating the monetary costs.

Income-based Value of Time Approach

For the income-based VOT measure, we follow the conventions in the literature and convert self-reported annual income to an hourly wage metric. To estimate the hourly wages per individual, we divide income by the annual hours worked per individual.[38] As part of the survey, respondents reported their average weekly hours spent working and we multiply this number by the number of working weeks per year to derive annual number of hours worked.[39] We assumed a total of 51 working weeks (2,040 hours per year with 40 hrs per week) as this is the most common assumption on annual hours in the literature (Parsons, 2003). For non-workers (retired and unemployed), the wage rate is assumed to be equal to zero. To derive $\text{VOT}_{1/3\text{wage}}$, we divide the imputed hourly wage by 3.

Individual-Specific Value of Time Approach

The individual-specific VOT approach consists of three steps. First, we estimate the VOT_{ISS} for the three summer months (June, July, August). Second, we derive the relative values of leisure time for the other two time periods (Winter and Fall). Third, we use the relative value of leisure time estimates to weigh the VOT_{ISS} estimates from the first step to yield three VOT_{ISS} estimates per person: Winter, June/Summer, and Fall.

Step 1: Estimate VOT in the three summer months (June, July and August)

The VOT_{ISS} in the three summer months (June, July and August) is estimated using

[38] Of the 823 respondents who completed the survey, 52 (6%) did not answer the income question. Missing data on socio-demographic information is imputed using the multiple imputation using chained equations (MICE) technique and all the other socio-demographic data.

[39] We bound the annual number of hours using a lower bound of 300 and an upper bound of 4,000 and setting individuals hours worked equal to these bounds if they report hours worked below or above them. A total of 9 people reported annual hours worked outside these bounds.

responses to the time valuation questions.[40] There are two WTA question formats included in the survey: a focus group context (presented in Figure 2.B.2) and a short-term work contract setting where respondents were asked to work for a company sorting paper files alphabetically. To assess whether the respondents understood the WTA question, all responses were checked to examine whether the probability of accepting the payment amounts increased as the payment levels increased. Of the 823 respondents, 23 respondents made errors in the focus group question, 10 made errors in the work contract question, and 2 made errors in both questions.[41] These responses are excluded from the analysis. The polychotomous responses for each payment level are converted into a binary variable using the 'Probably Yes (75%)' as the lower bound cut-off for a 'yes' response. This procedure yields a total of 4,903 choices from 789 individuals for analysis.

A random parameters logit model is specified to accommodate observed and unobserved heterogeneity and to derive individual-specific VOT estimates (Hensher et al., 2015). The payment amount is included as a fixed parameter and two random parameters are included: a constant for giving up the 8 hours and a dummy variable for the work contract question. To illustrate the random parameters model, we assume that individual q faces a choice among J alternatives in each of M choice situations. Thus, individual's q utility associated with each alternative j in each choice situation m can be represented as:

$$U_{jmq} = \beta_p p_{jm} + \beta_q' x_{jm} + \epsilon_{jmq}$$

where p_{jm} is the payment amount associated with alternative j in choice situation m, β_p is a fixed coefficient to be estimated, x_{jm} is a vector of attributes of the alternatives, β_q is a vector of random parameters that are assumed to vary across individuals, and ϵ_{jmq} is an independent and identically distributed (IID) extreme value type 1 error term.

Observed individual heterogeneity is introduced into the utility function through β_q. In our specification, socio-demographic characteristics are incorporated as affecting the means of the two random parameters, and thus we write

$$\beta_q = \beta + \Delta Z_q + \eta_q$$

[40] Although a willingness-to-pay (WTP) format time-money trade-off question was asked of working respondents where they had the opportunity to buy time off from their employers, in this study we focus on the WTA questions because these contexts better reflect the fact that individuals are giving up their leisure time to participate in recreational fishing and the WTP questions were not asked of non-working respondents.

[41] To assess possible ordering effects of the specific time questions as well as whether the time questions are asked to respondents before or after the fishing behaviour questions, ordering effect dummies are included in the model. In both cases, the ordering effects dummy variable is insignificant.

where Z_q are the socio-demographic characteristics, Δ is a matrix of coefficients to be estimated, and η_q is an individual-specific random component. Specific socio-demographic characteristics included in the model are employment status (working full or part-time, self-employed, or not working), a dummy variable whether the respondent is male, education level (some college or less, a bachelor's/associate's degree, or a graduate degree), a dummy variable if the respondent's household income is above $100,000, household size, a dummy variable whether the respondent has children, and the age index variable.

We tested several different potential distributions for the random parameters but ultimately settled on the scaled beta distribution because it restricts the coefficients to be positive (i.e. avoids negative individual WTA estimates), avoids the issue with unrestricted distributions such as log-normal of giving implausible results for some share of the population, and yields a smoother distribution compared to the triangle or uniform distributions (Hensher et al., 2015).[42] Table 2.C.1 presents the results for the time valuation random parameters logit model. The payment amount has the expected positive impact on the likelihood of giving up the 8 hours of time. The presence of observed heterogeneity is illustrated by the significance of many of the socio-demographic characteristic coefficients. The coefficients for the random parameters themselves are more complicated to interpret because of the multiple socio-demographic interaction terms.[43]

Using the random parameters logit model estimates, the VOT can be calculated as the ratio of the expected value of β_q divided by β_p. To derive individual-specific estimates of the VOT, we utilize additional information about the choices each individual makes as well as socio-demographic characteristics to compute conditional distributions for each individual. The VOT_{ISS} estimates are calculated as the mean of these conditional distributions.[44] We use the focus group VOT_{ISS} estimates which are slightly lower than the work contract estimates. The VOT_{ISS} estimates for each individual are converted to an hourly basis. For respondents with missing VOT estimates either because of a misunderstanding error or they did not complete the questions, values are calculated by multiple imputation using chained equations (MICE) technique and the same set of socio-demographic variables.

Step 2: Estimate Marginal Rates of Substitution for Leisure Days

[42] The specific form of the scaled beta distribution is $\beta_q = \beta\nu_q,\ \nu_q \sim beta(3,3)$.

[43] The constant and work contract were multiplied by -1 before estimation to ensure the coefficient is negative (Hensher et al., 2015).

[44] Details on the specific steps to compute individual-specific estimates from random parameters logit models is presented in Hensher et al. (2015).

Table 2.C.1: Random parameters logit model estimates for parameters of value of summer time

Variable	Estimate	
Nonrandom parameters		
Payment amount ($00s)[a]	0.890***	
	(0.026)	
Random parameters (Scaled beta distribution[b])		
Constant[c]	-2.815***	
	(0.425)	
Work contract[c]	1.693***	
	(0.612)	
Heterogeneity in means of random parameters		
	Constant	Work Contract
Work full- or part-time	1.888***	-2.757***
	(0.153)	(0.231)
Self-employed	-0.393	-0.228
	(0.254)	(0.375)
Income >$100,000	-0.321**	0.053
	(0.135)	(0.204)
Male	-0.236	-1.287***
	(0.173)	(0.259)
Bachelor degree	-0.119	-0.049
	(0.142)	(0.209)
Graduate degree	-1.843***	0.123
	(0.181)	(0.280)
Ageindex	-3.269***	-1.425***
	(0.280)	(0.408)
Household size	-0.148**	0.459***
	(0.067)	(0.101)
Children	0.123	-0.726***
	(0.188)	(0.277)
Number of observations	4,903	
Number of respondents	789	
Log-likelihood	-2,322.4	

Notes: Standard errors in parentheses. ***, **, * are significance at 1%, 5%, 10% level.
[a] Payment amount was rescaled to hundreds of dollars for computational reasons.
[b] The scaled beta distribution is represented as $\beta_q = \beta\nu_q,\ \nu_q \sim beta(3,3)$.
[c] Variable was multiplied by -1 prior to estimation to ensure the coefficient is negative, and the parameter estimates in the table have been multiplied by -1 again for interpretability.

The VOT_{ISS} estimates from the previous section are for the three summer months. For the other time periods of the year, we need to scale the VOT in the summer months by the relative value of leisure days in the other two time periods. The relative value is calculated as the marginal rates of substitution (MRS) for leisure days and represents how many summer days an individual would be willing to give up to obtain one fall or winter day. For all MRS calculations, summer months are represented as the 'numeraire' and thus we calculate two metrics: the MRS of fall days for summer days ($MRS_{f,s}$) and the MRS of winter days for summer days ($MRS_{w,s}$). As an illustration, a $MRS_{f,s}$ estimate of 0.5 would suggest that a respondent is willing to give up 2 fall leisure days to obtain one summer leisure day.

We use the responses to the leisure day trade-off questions presented in Figure 2.B.3 of to calculate the MRS of leisure days. Because a significant proportion of respondents (360 out of 823 respondents) chose the option with no changes in their leisure days (Option C) as their most preferred for both leisure trade-off questions, we decided to treat these respondents separately. For these respondents, we know that $MRS_{f,s}$ and $MRS_{w,s}$ are bounded from above by one divided by the maximum number of fall and winter leisure days that the respondent could have gained by giving up one summer day. For these cases, we set the MRS to be equal to this upper bound to yield a more conservative estimate of the MRS (i.e. an MRS closer to one). For example, if a respondent chose the status quo option but could have gained 3 fall days or 2 winter days for giving up one summer day, then $MRS_{f,s}$ is calculated to be 0.33 and $MRS_{w,s}$ is calculated to be 0.5.

For the remaining respondents, we estimate a random parameters logit model and derive individual-specific estimates of MRS conditional on their choices and individual characteristics using the same approach as described in the Step 2.[45] Both the number of fall and winter days are included as random parameters, but summer days, as our numeraire, is included as a fixed parameter. The same set of individual characteristics as the VOT models are included in the model as well as three additional time flexibility variables: Fixed hourly schedule is a dummy variable for respondents who are not free to set work schedule, Prefer to work less is a dummy variable for respondents who stated they would prefer to work less, and Prefer to work more is a dummy variable for respondents who stated they would prefer to work more. Socio-demographic characteristics are included as affecting the means of fall and winter days, but not summer days. We considered

[45]Because the leisure trade-off questions asked for both the respondent's most and least preferred option amongst the three presented, we include both types of responses by switching the sign of the attributes for the least preferred options. Implicitly, this approach assumes a symmetry of preferences between most/least preferred. Estimating separate models for the most and least preferred choice subsamples results in similar estimates of the MRS.

Table 2.C.2: Random parameters logit model estimates for parameters of marginal rates of substitution of leisure

Variable	Estimate	
Nonrandom parameters		
Summer	1.365***	
	(0.136)	
Random parameters (Scaled beta distribution[a])		
Fall	1.101***	
	(0.249)	
Winter	3.119***	
	(0.253)	
Heterogeneity in means of random parameters		
	Fall	Winter
Work full- or part-time	0.867***	-0.709***
	(0.128)	(0.131)
Self-employed	0.931***	-0.504***
	(0.152)	(0.152)
Fixed hourly schedule	-0.647***	-0.448***
	(0.095)	(0.094)
Prefer to work less	0.161*	0.688***
	(0.097)	(0.097)
Prefer to work more	0.179	-0.382***
	(0.122)	(0.112)
Income >$100,000	-0.031	0.002
	(0.079)	(0.079)
Male	0.490***	0.411***
	(0.094)	(0.097)
Bachelor degree	0.002	0.509***
	(0.078)	(0.079)
Graduate degree	0.425***	0.013
	(0.110)	(0.106)
Ageindex	0.588***	-0.532***
	(0.163)	(0.165)
Household size	-0.110***	-0.116***
	(0.037)	(0.036)
Children	0.078	0.140
	(0.095)	(0.096)
Number of observations	1,842	
Number of respondents	463	
Log-likelihood	-1,822.3	

Notes: Standard errors in parentheses. ***, **, * are significance at 1%, 5%, 10% level.
[a] The scaled beta distribution is represented as $\beta_q = \beta\nu_q,\ \nu_q \sim beta(3,3)$.

several different distribution forms for the random parameters and ultimately settled on the scaled beta distribution for reasons discussed in Step 2. Table 2.C.2 presents the results for the MRS model.

Step 3: Derive Value of Time Estimates for Fall and Winter

As a final step to yield the $\mathrm{VOT}_{\mathrm{ISS}}$ in the other time periods, we multiply the $\mathrm{VOT}_{\mathrm{ISS}}$ in summer derived in the Step 1 by $MRS_{f,s}$ and $MRS_{w,s}$ derived in Step 2 to calculate the $\mathrm{VOT}_{\mathrm{ISS}}$ in fall and winter.[46]

Travel Costs

We combine the VOT estimates derived using the two approaches in the previous sections with the monetary costs of traveling to each port and the fees to take a headboat fishing trip. The monetary travel costs are calculated generally following the approach described in the Deep Water Horizon Recreation Assessment Study's (DHW study) technical memoranda documents (Industrial Economics Inc, 2015; Leggett, 2015). We consider that individuals can either fly or drive to each port and travel costs are calculated as a weighted average of these costs where the weights are the probabilities of choosing each mode of travel. Formally, the cost C^{TC}_{iojt} to individual i incurred from traveling from origin o to port j in time period t is represented as

$$C^{TC}_{iojt} = \pi_{ioj} C^{Fly}_{iojt} + (1 - \pi_{ioj}) C^{Drive}_{iojt}$$

where π_{ioj} represents the probability that individual i will choose to fly when traveling from origin o to port j. Origin locations are assigned by geocoding the zip code provided by respondents in the survey. Time period t refers to one of the four seasonal time periods in each of the two years.

The Costs of Driving

Driving costs are calculated using information on both monetary and non-monetary expenses. Data on average per-mile fuel costs (f_t) and average per-mile non-fuel vehicle operation costs (nf_t) including tires, maintenance and depreciation is collected from the AAA (American Automobile Association, 2015; Association, 2016).[47] One-way driving

[46]Note that although we have four fishing time periods, we only have three $\mathrm{VOT}_{\mathrm{ISS}}$ estimates per individual as we assume that the $\mathrm{VOT}_{\mathrm{ISS}}$ in June is equal to time in July and August.

[47]Costs are obtained for an average sedan and depreciation costs are calculated using 5,000 mile deviations (higher and lower) from the 15,000-mile annual depreciation scenario reported by the AAA following the DWH approach. Specifically, the AAA estimates that in 2014 depreciation costs for the average sedan are \$252 less if the car is driven 5,000 miles less than the 15,000-mile scenario and \$204 more if the car

distances (in miles) and time (in hours) between any given points a and b are estimated from Google Maps using the R package ggmaps (Kahle and Wickham, 2013). The average cost of a night at a hotel (h_t) is obtained from the American Hotel and Lodging Association (American Hotel & Lodging Association, 2015). The number of nights is derived by dividing total driving time by 12 and rounding down to the nearest integer. These monetary costs of driving are divided by the fishing party size reported by each individual (ρ_i) in the survey. If the reported party size is greater than 5, then the costs are divided by 5 to reflect the capacity of a typical sedan.[48] Thus, the one-way driving costs is calculated as,

$$C_{it}(a,b) = \left[(f_t + nf_t) * distance(a,b) + h_t * nights(a,b)\right]/\rho_i + VOT_{it} * time(a,b),$$

where VOT_{it} is the value of leisure time for individual i in time period t. These one-way costs are multiplied by two to derive the round-trip cost to each individual.

$$C^{Drive}_{iojt} = 2 * C_{it}(origin_{io}, port_j)$$

The Costs of Flying

To calculate the costs of flying, we first identify several possible flying routes for each individual and then choose the flying route with the least cost. Specifically, the four closest origin airports m to each individuals' residences are identified along with the four closest destination airports n to each visited port, for a total of 16 potential flying routes for each individual to each port.[49] The costs of flying is then estimated to be the minimum cost route amongst these possible pairs:

$$C^{Fly}_{iojt} = min_{m,n}\left\{C^{Fly}_{iojtmn}\right\}$$

The costs of flying can be divided into five parts: the costs of driving from the origin location o to the origin airport m, the costs of parking at the origin airport, the flight costs from the origin airport to destination airport n near the port j, the cost of renting a car, and the cost of driving from destination airport to the port (Leggett, 2015). These

is driven an additional 5,000 miles. The average per-mile depreciation costs is 0.0511((252/5,000) + ($204/5,000))/2).

[48]If the party size information was missing, then the median party size of 3 is used in its place.

[49]We only consider airports that have annual enplanements of over 100,000 (Industrial Economics Inc, 2015). Enplanement data is obtained from the Federal Aviation Administration (FAA) Calendar Year 2014 Passenger Boarding and All-Cargo Data lists.

different components are represented using the following expression:

$$C^{Fly}_{iojtmn} = 2 * C_{it}(origin_{io}, airport_{iom}) + C^{Parking}_{mt} + C^{Flight}_{itmn} + C^{Rental}_{t} + 2 * C_{it}(airport_{jn}, port_j)$$

Both driving portions of the costs are calculated using the same methodology as the cost of driving directly to the port as described previously. Average parking costs for large/medium and small airports are based on data used in the DHW study. The number of parking days is calculated using the number of total nights away reported by respondents and all parking costs are weighted by the reported party sizes. Average rental car costs are based on data reported by the American Hotel and Lodging Association (American Hotel & Lodging Association, 2015).

Total round-trip flying costs from origin airport n to destination airport m is calculated as

$$C^{Flight}_{itmn} = VOT_{it}(time^{airport} + time^{flight}_{tmn} + time^{layover}_{tmn}) + price^{ticket}_{tmn}$$

where VOT_{it} is the value of time, $time^{airport}$ is the time spent at the origin and destination airports before and after the flights and is assumed to be 4 hours for each round-trip, $time^{flight}_{tmn}$ is the flight time between airports, $time^{layover}_{tmn}$ is the time spent during any layovers, if any, and $price^{ticket}_{tmn}$ is the round-trip ticket price.

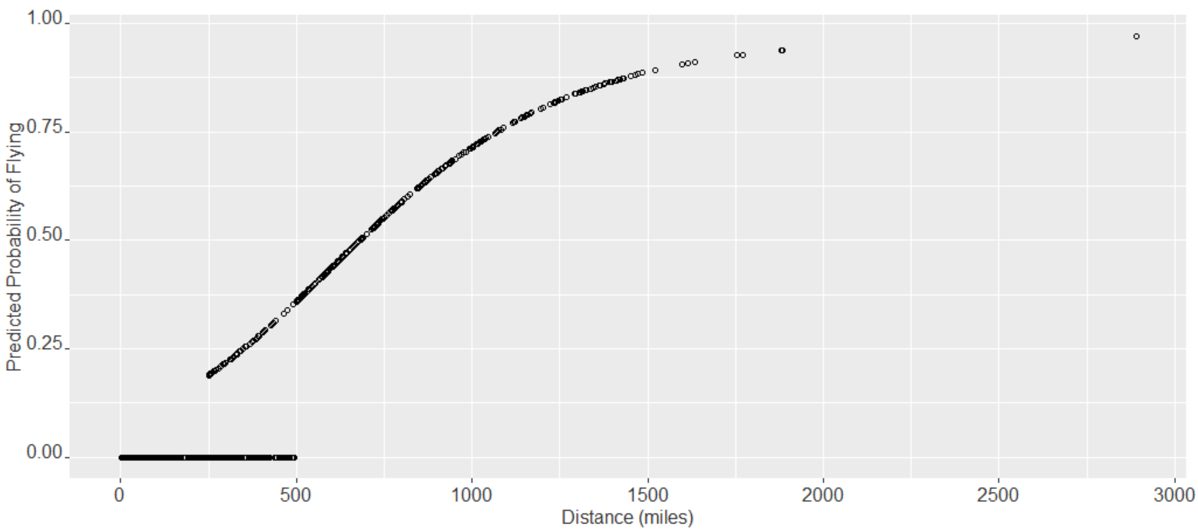
These last three terms are based on data from the Airline Origin and Destination Survey (DB1B) conducted by the Office of Airline Information of the Bureau of Transportation Statistics. The DB1B dataset represents a 10% sample of airline tickets from reporting carriers in the United States every year. The DB1B consists of three different data tables (*ticket*, *market*, and *coupon*) that can be joined by an Itinerary ID variable. The *ticket* data table is the most aggregated and includes information on ticket fare costs and the origin of the flight but does not include destination. It represents each purchased ticket by one observation (either one-way or round-trip). The destination information is found in the *market* data table which includes one observation per direction of travel (i.e. one observation for a one-way ticket and two observations for a round-trip ticket). The *coupon* data table includes one observation per flight of the journey and has information on estimated and actual flight time per leg as well as the number of layovers. We obtain data for each quarter of 2014 and 2015.

For each origin-destination airport pair, the average number of layovers and flight time are obtained from the *coupon* database for each quarter of 2014 and 2015. Some flight routes had missing time data for certain time periods. In these cases, a regression model is estimated using distance (in miles) and number of layovers as explanatory variables to predict the expected flight times for a small subset of routes. For each layover, 60 additional minutes are added to the total flight time based on data from Sabre Airline Solutions (Industrial Economics Inc, 2015). The ticket fares are taken from the *ticket* data table and the 30th percentile fare is used as the expected flight costs (Industrial Economics Inc, 2015). Finally, the average flight times and costs for each quarter are converted to our four seasons.

Probability of Flying

The probability of flying is modeled as a function of distance using data on actual mode choices from the 2009 National Household Travel Survey (NHTS). The NHTS survey collects information on mode of transportation and distance for a nationally representative sample of travel behaviour. After excluding trips that are less than 250 miles, we are left with 2,393 trip decisions that are used in a logit model of the decision to fly or drive based on miles and miles squared. The estimated intercept and distance parameters are used to predict the probability of flying for respondents in our data. Following DHW study, we assign a zero probability of flying to all respondents who reside less than 250 miles away from the port and to respondents who live less than 500 miles away and have income less than $70,000 per year or more than 2 household members. Figure 2.C.1 shows the predicted probability of flying for respondents in our data based on their distance to the port.

Figure 2.C.1: Predicted probability of flying for individuals as a function of distance



Multiple Ports

The vast majority of respondents only visited one port. Respondents visiting more than one port, the different ports are usually close by. For individuals that traveled to more than one port, we average the travel costs over visited ports to derive an average cost of a trip to the Gulf of Mexico.

Second Homes

For respondents reporting a second home in the region, we also calculate the expected travel costs from their second home to each visited port. A total of 53 respondents reported second homes in the region with valid zip codes. We used the travel costs from the second home if these expected costs are lower than the travel costs estimated from the main residence.

Total Costs of a Fishing Trip

In addition to the costs of traveling to the port, headboat fishermen also pay a trip fee to go fishing. Thus, the total costs of a headboat fishing trip for individual i is comprised of the costs of traveling from origin o to port j and the costs of a fishing trip of length l from port j in time period t:

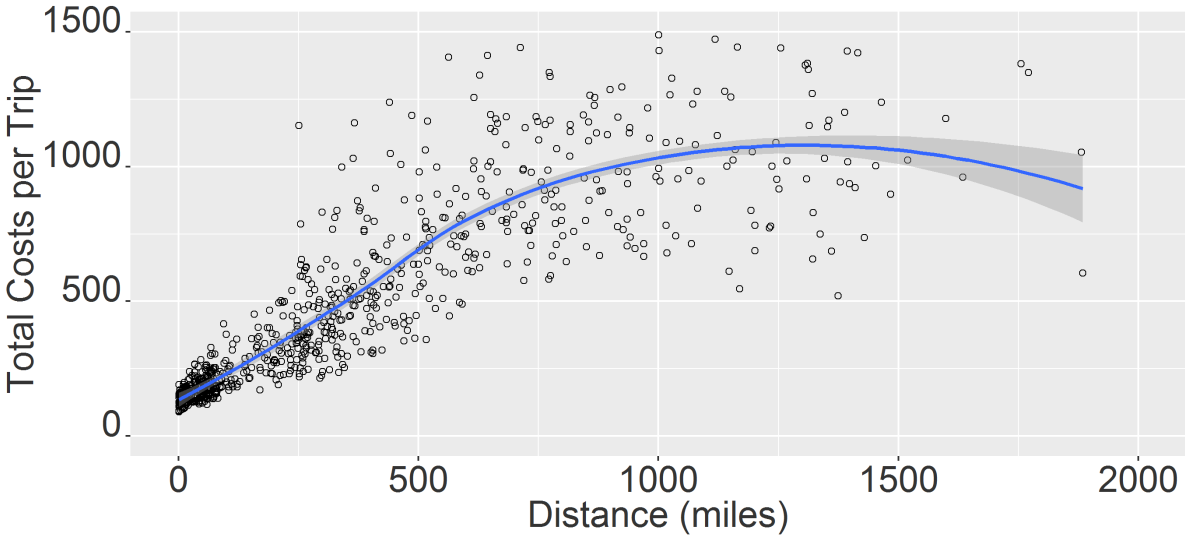
$$C_{ijtl} = C_{iojt}^{TC} + C_{jtl}^{Fish}$$

These fees typically depend on the specific port and the length of the fishing trip but may also vary seasonally. For the contingent behaviour data, we include the trip price as presented to respondents in the survey. These trip costs ranged from \$50 to \$150 for partial day trips and \$80 to \$250 for full day trips. For the revealed preference data, we collect information on the rates charged by headboat operators in 2014 and 2015 at each port from an online survey of headboat operators. The average trip fees for headboat trips in each port ranged from \$50 to \$130 for partial day trips to \$80 to \$250 for full day trips. For ports with missing trip fee information, we used the closest port's trip fee data in its place.

Travel Cost Summary

Figure 2.C.2 shows the relationship between the total costs per trip and distance for all individuals using the VOT_{ISS} estimates. As shown, travel costs have a nonlinear relationship with distance due to the high fixed costs of flying.

Figure 2.C.2: Relationship between total travel costs per trip and distance



Appendix 2.D: Value of time robustness checks

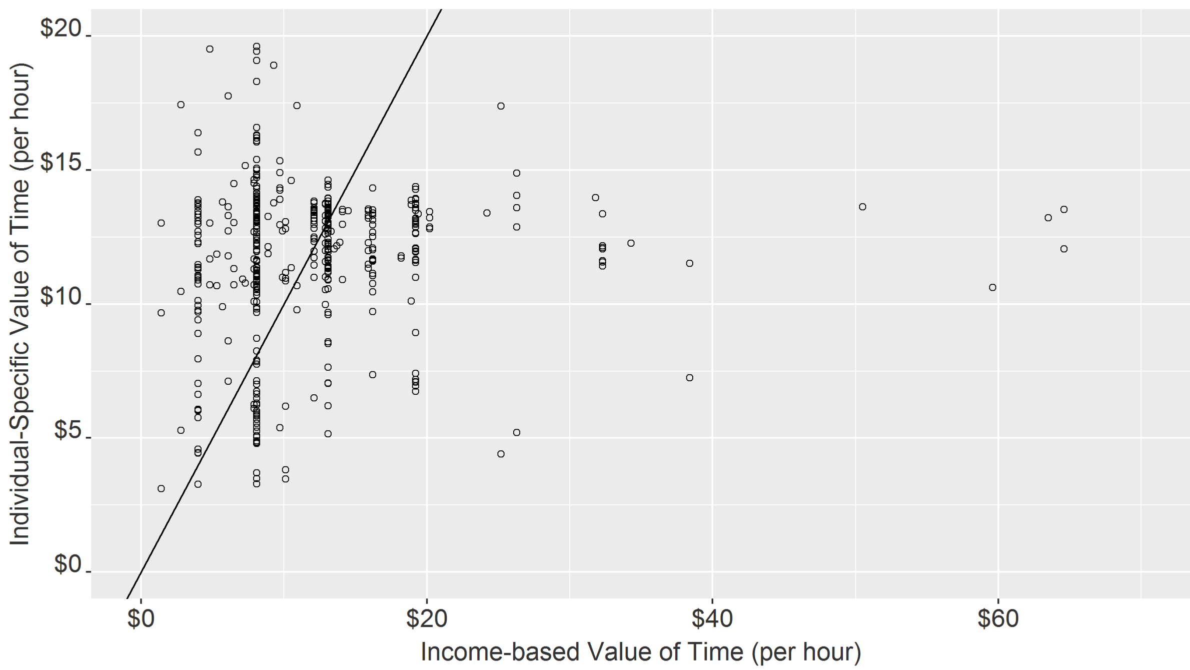
To check the validity of the individual-specific approach to time valuation, we collect three pieces of supporting evidence. We check our model specification, replicate the VOT comparison to the income-based measure in another context, and we investigate whether value estimates from this approach reflect actual choices made in real contexts. First, we employ alternative model specifications for the individual-specific approach to check the robustness of our results. We examine using a latent class specification and alternative distributions for random parameters instead of beta (including Weibull, triangle, normal, gamma, and exponential). We also used the VOT associated with the work contract format question which was generally higher than the focus group question. Finally, we used various subsets of the socio-demographic characteristics. Across these different specifications, the VOT varies from \$21 to \$35. More importantly, all model specifications produce VOT estimates that have a similar low correlation with the income-based measure.

We use the data on road toll and driving time decisions to beaches in Italy in Fezzi et al. (2014) to provide an alternative income-based VOT measure comparison using revealed preference information. We estimate a random parameters logit model using a similar approach to the current study as outlined in Appendix 2.C. We derive individual-specific VOT estimates and compare these estimates to hourly wages derived from income reported in the data. Figure 2.D.1 presents the same scatterplot as Figure 2.1 but using the data in Fezzi et al. (2014). We find that the individual-specific VOT estimates are

50% to 70% of the wage, which is similar to the magnitudes found in the current study. Furthermore, the correlation coefficient between the two VOT measures is 0.11. Thus, the revealed VOT from a road toll context also has a similar low correlation with the income-based approach.

A final check on our individual-specific approach is to assess whether the resulting VOT estimates reflect real decisions involving money. In other work, we conduct laboratory experiments using the same SPC format where respondents are asked to give up their time to help in the library in exchange for monetary payments (Lloyd-Smith and Adamowicz, 2017). These same individuals were then given the opportunity to make a real, binding monetary offer for how much they would need to be paid to work in the library. We find that the VOT estimates derived from the same time valuation choice questions are similar to the real offers made by participants. Specifically, the average VOT estimate derived from the time valuation questions is \$17.35 (median \$15.00) compared to an average real offer of \$18.00 (median \$15.00).[50] The overall correlation between the stated and real values is 0.62.

Figure 2.D.1: Relationship between value of time estimates using individual-specific and income-based approaches in Fezzi et al. (2014)



[50]Additional information on the experiment and results is provided in Lloyd-Smith and Adamowicz (2017).

Appendix 2.E: Robustness checks on welfare results

To decompose the role the different parameter estimates and VOT measures in the welfare calculations, we simulate the welfare measures using Model 1's parameter estimates and $VOT_{1/3wage}$ measures and with Model 2's parameter estimates and VOT_{ISS} measures. Table 2.E.1 summarizes the welfare measures using the various combinations of parameter estimates and VOT measures.

Table 2.E.1: Welfare estimates

Parameter Estimates	Model 1	Model 2	Model 1	Model 2
VOT Measure	VOT_{ISS}	$VOT_{1/3wage}$	$VOT_{1/3wage}$	VOT_{ISS}
Policy Scenario	Mean welfare impacts (person/year)			
Policy 1: Trip fee increase all year	-$124	-$119	-$119	-$125
($25 partial/$50 full day)	(0.21)	(0.28)	(0.27)	(0.24)
Policy 2: Summer fishing closure	-$193	-$116	-$115	-$196
	(4.4)	(2.7)	(2.6)	(4.6)
Policy 3: Fall fishing closure	-$216	-$144	-$143	-$218
	(5.1)	(3.5)	(3.3)	(5.3)

Notes: Cluster bootstrap standard errors in parentheses.
Estimates are generated with 500 conditional error draws per individual.

Appendix 2.F: A new approach to calculating welfare measures in Kuhn-Tucker demand models

In this appendix I develop a new approach to calculating welfare measures in Kuhn-Tucker consumer demand models that uses the analytical properties of the Multiple Discrete-Continuous Extreme Value (MDCEV) specification. I adapt Pinjari and Bhat's (2011) Marshallian demand forecasting routine to calculate Hicksian demands that are useful for computing welfare measures. Simulations demonstrate that this new approach substantially reduces computational time relative to the existing approach using a numerical bisection routine. The new approach performs best relative to the numerical bisection routine if i) a γ-profile utility function is specified, ii) the number of choice alternatives available is high, or iii) the average number of chosen alternatives is low.

Introduction

Many individual choice contexts can be characterized by both extensive (i.e. what alternative to choose) as well as intensive (i.e. how much of an alternative to consume) margins where individuals are not restricted to only choosing a single alternative. These multiple discrete-continuous (MDC) choice situations are ubiquitous, arising in transportation, marketing, and decisions regarding environmental resources.[51] Kuhn-Tucker (KT) consumer demand models are often employed to analyze these MDC situations and substantial progress has been made on improving the econometric modeling structures. One reason cited for the lack of widespread use of these KT models is that applying these models for welfare analysis is not straightforward (von Haefen and Phaneuf, 2005; Bhat and Pinjari, 2014). This issue is especially relevant in applying these models to studying decisions regarding environmental resources where producing welfare estimates is often the main purpose of the research.

Computing exact welfare measures from individual demand models face many difficulties.[52] While the theoretically correct welfare measures are based on Hicksian demands, which hold utility constant and can be used to compute compensating and equivalent variation, analysts often use demand models that provide estimates of Marshallian demands and their associated consumer surplus welfare measure (Bockstael and McConnell,

[51] Bhat and Pinjari (2014) review the MDC choice model literature and discuss relevant empirical applications.

[52] For a textbook treatment of these issues, see Bockstael and McConnell (2007).

2007; Laird, 2010). A large theoretical and empirical literature has focused on the appropriateness of calculating Hicksian welfare measures from Marshallian demands which has motivated the development of a diverse set of approaches that use various approximations and assumptions on the structure of equations (Willig, 1976a; Jara-Diaz and Videla, 1990; Bockstael and McConnell, 2007; Dalya et al., 2008; Laird, 2010). Overall the literature provides mixed evidence on whether consumer surplus is a good proxy for Hicksian welfare measures. Calculating Hicksian demands directly avoids the approximations and assumptions that are required when starting with Marshallian demands. One of the advantages of the KT modeling framework is that the utility function is explicitly specified, which allows for the direct computation of exact Hicksian welfare measures. This fact motivates the approach developed in this paper.

This paper describes a new approach to calculating Hicksian welfare measures in KT consumer demand models. The main difficulty in calculating the optimal consumption quantities for individuals in KT models is that once the model parameters are estimated, a constrained, non-linear optimization problem needs to be solved. The existing iterative approach, using a numerical bisection routine (von Haefen, 2007), works well in most applications and is computationally more efficient than earlier enumerative approaches, where every possible solution is checked (Phaneuf et al., 2000). However, its iterative nature is undesirable in more data intensive applications and relies on the arbitrary choice of a stopping criteria. The new approach presented here uses analytical properties and expressions of the Multiple Discrete-Continuous Extreme Value (MDCEV) utility specification (Bhat, 2008) to significantly reduce computation time. I adapt Pinjari and Bhat's (2011) Marshallian demand forecasting routine to calculate Hicksian demands that can be used to compute exact welfare measures. Simulations using a real data set suggest that using the new algorithms can reduce computation time 3- to 12-fold compared to the existing iterative approach. Experiments conducted using simulated data also demonstrate that the new approach's relative computational performance is best when the number of choice alternatives available is high or the average number of chosen alternatives is low.

The individual's expenditure minimization problem

I start by considering the general MDCEV utility function as in Bhat (2008):

$$U(x) = \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left[\left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right] + \frac{\psi_1}{\alpha_1} x_1^{\alpha_1}. \tag{2.F-1}$$

where x_k is the amount of K alternatives available to the decision maker and x_1 is the numeraire or "outside" good that is always consumed in positive quantities. To be consistent with the properties of a utility function $\gamma_k > 0$, $\psi_k > 0$ and α_k, $\alpha_1 \leq 1$ for all k are required for this function (Bhat, 2008). Although the standard assumption is to assume the price of the numeraire is equal to one, I use p_1 throughout this paper for clarity. The ψ_k, γ_k, and α's terms are structural parameters of the utility function and Bhat (2008) provides a thorough overview of the interpretation of these parameters. In brief, ψ_k is the marginal utility of alternative k when $x_k = 0$, the α's are satiation parameters and control the rate of diminishing marginal utility of additional consumption of an alternative, and γ_k shifts the underlying indifference curves which allows for corner solutions (i.e. zero consumption levels for certain alternatives).

Individuals are assumed to maximize utility given by Equation (2.F-1) subject to a linear budget constraint and non-negativity constraints on x_k. Pinjari and Bhat (2011) solve this consumer problem to yield analytical expressions for Marshallian demands. However, for welfare analysis, we are interested in Hicksian demands and thus I set up the consumer's expenditure minimization problem holding utility constant at the baseline level ($\bar{U}$). Specifically, the consumer's expenditure minimization problem is

$$\min_{\sum_{k=1}^{K} x_k} \quad E = \sum_{k=1}^{K} p_k x_k \ \text{ subject to } \ U(x) = \bar{U}$$

and non-negativity constraints on the Hicksian demand consumption, x_k. The Lagrangian equation is then given by

$$L = \sum_{k=1}^{K} p_k x_k + \lambda^E \left[\bar{U} - \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left[\left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right] - \frac{\psi_1}{\alpha_1} (x_1)^{\alpha_1} \right],$$

where λ^E is the Lagrangian multiplier associated with the baseline utility constraint. The resulting KT first-order conditions for optimal expenditures are given by:

$$\frac{\partial L}{\partial x_1} = p_1 - \lambda^E \psi_1 (x_1)^{\alpha_1 - 1} = 0 \tag{2.F-2}$$

$$\frac{\partial L}{\partial x_k} = p_k - \lambda^E \psi_k \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k - 1} = 0, \ \text{ if } \ x_k > 0, \ \ k = 2, ...K, \tag{2.F-3}$$

$$\frac{\partial L}{\partial x_k} = p_k - \lambda^E \psi_k \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k - 1} > 0, \ \text{ if } \ x_k = 0, \ \ k = 2, ...K, \tag{2.F-4}$$

$$\frac{\partial L}{\partial \lambda^E} = \bar{U} - \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left[\left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right] - \frac{\psi_1}{\alpha_1} x_1^{\alpha_1} = 0. \tag{2.F-5}$$

These first-order conditions can be used to derive Hicksian demands and welfare measures. The Hicksian compensating surplus (CS^H) for a change in price and quality from baseline levels p^0 and q^0 to new 'policy' levels p^1 and q^1 is defined implicitly using an indirect utility function

$$V(p^0, q^0, y, \theta, \varepsilon) = V(p^1, q^1, y - CS^H, \theta, \varepsilon) \tag{2.F-6}$$

or explicitly using an expenditure function

$$CS^H = y - e(p^1, q^1, \bar{U}, \theta, \varepsilon) \tag{2.F-7}$$

where y is income, θ is the vector of structural parameters $(\psi_k, \alpha_k, \alpha_1, \gamma_k)$, ε is a vector or matrix of unobserved heterogeneity, and $\bar{U} = V(p^0, q^0, y, \theta, \varepsilon)$. Before describing the new approach to welfare measurement, I first review main difficulties in welfare measurement in KT models and describe the existing approaches employed in the literature.

Calculating welfare measures in Kuhn-Tucker models

There are two main computational difficulties in solving for CS^H using either Equation (2.F-6) or (2.F-7) (von Haefen and Phaneuf, 2005). First, the unobserved heterogeneity term is unknown to the analyst and therefore CS^H is a random variable from the analyst's perspective. Second, a priori, the analyst does not know which non-numeraire alternatives have a positive consumption level (i.e. interior solution) or are not consumed (i.e. corner solutions). Thus, Equations (2.F-6) and (2.F-7) can be viewed as endogenous regime switching functions where there are 2^K-1 possible combinations of interior and corner solutions for the non-numeraire alternatives.

To address the first issue, Monte Carlo simulation techniques can be employed to draw simulated values of the unobserved heterogeneity (ε) and then a measure of the central tendency of CS^H such as its expectation can be computed. The main implication of using these Monte Carlo techniques is that now CS^H needs to be computed for each simulated value of ε which can be burdensome if the approach used to calculate CS^H is slow. This implication has partially motivated the use of conditional draws of ε which avoids the need to calculate consumption patterns in the baseline state as well as simulate the entire distribution of unobserved heterogeneity. The conditional approach to welfare measurement draws ε such that the model perfectly predicts the observed consumption patterns in the baseline state (von Haefen and Phaneuf, 2005).

With the error draws in hand, the second step to welfare measurement in KT models is to solve for CS^H. There have been three approaches employed in the literature to calculate CS^H conditional on the simulated unobserved heterogeneity. Each of these approaches have iteratively refined the previous approach to improve the computational performance. Phaneuf et al. (2000) introduced the modern approach to welfare measurement in KT models and proposed an enumerative approach that considered all the possible combinations of alternatives that the decision-maker could choose to consume. To ensure the utility for each individual is the same in the baseline and policy states, a numerical bisection routine is needed to iteratively solve for income. For a given level of income at each iteration of the numerical bisection routine, this brute force method solved for all $2^K - 1$ indirect utility functions for each draw of ε and then determined which one has the highest associated utility. While this approach computes consistent estimates of CS^H, the two main practical limitations of this approach is that it becomes intractable as K increases and the preference specification must have a closed form conditional indirect utility function.[53]

von Haefen et al. (2004) refine this earlier approach by replacing the enumerative procedure with an iterative algorithm that uses the KT conditions to solve for an individual's optimal consumption bundle and utility. In addition to reducing the computational time, this approach can be used with preference specifications that do not have closed form indirect utility functions such as the MDCEV structure in Equation (2.F-1) or the commonly used specifications in environmental economics described in von Haefen and Phaneuf (2005). However, this approach still uses the implicit definition of CS^H shown in Equation (2.F-6) and therefore still requires the use of a higher level numerical bisection routine to solve for income by equating utilities in the baseline and policy states.

von Haefen (2007) builds on von Haefen et al. (2004) using a more efficient numerical approach by making use of the explicit definition of CS^H shown in Equation (2.F-7). The approach exploits the additively separable utility function structure and replaces the two level numerical bisection routines in von Haefen et al. (2004) with a single numerical bisection routine. I fully describe the steps of the algorithm below because I use this approach as a benchmark to compare the computational performance of the new approach introduced in this paper. For the rest of the paper, I refer to this approach as the numerical bisection algorithm.

[53] Phaneuf and Herriges (1999) demonstrate that this approach becomes infeasible with choice sets beyond around 15 non-numeraire alternatives.

The numerical bisection algorithm (von Haefen, 2007).

The logic of this approach is to note that if we know the optimal numeraire quantity, x_1, we could solve for λ^E and use this in Equation (2.F-3) to solve for x_k. Thus the following numerical bisection routine solves for x_1:

Step 0: Set $x_{1l}^0 = 0$ and $x_{1u}^0 = \frac{\alpha_1}{\psi_1}(\bar{U})^{1/\alpha_1}$ to initialize the algorithm.

Step 1: At iteration i, set $x_{1a}^i = (x_{1l}^{i-1} + x_{1u}^{i-1})/2$. Go to step 2.

Step 2: Plug x_{1a}^i into Equation (2.F-2) and solve for λ^E. Use λ^E in Equation (2.F-3) to solve for x_k^i. Solve for $U^i = U(x_k^i, x_{1a}^i)$ using Equation (2.F-5). Go to step 3.

Step 3: If $U^i < \bar{U}$, set $x_{1l}^i = x_{1a}^i$ and $x_{1u}^i = x_{1u}^{i-1}$. Else, set $x_{1l}^i = x_{1l}^{i-1}$ and $x_{1u}^i = x_{1a}^i$.

Step 4: Go back to step 1 and iterate until $|(x_{1l}^i - x_{1u}^i| \leq c$ where c is arbitrarily small.

The optimal Hicksian demands that result from this algorithm can be multiplied by the prices in the policy state to derive the expenditures necessary to reach baseline utility and CS^H can be solved using Equation (2.F-7).

A new approach to welfare calculation in Kuhn-Tucker models

This section introduces the new approach to computing welfare measures. I first derive the property that is used to discriminate between which non-numeraire alternatives are consumed and which ones are not consumed. Next, I introduce two versions of the algorithm that differ in the assumed utility function specification.

Differentiating between consumed and unconsumed alternatives.

The first step is to use the first-order conditions and the value of λ^E to show how to discriminate between which non-numeraire alternatives are consumed and which ones are not consumed. I introduce the notation i to denote non-numeraire alternatives with a positive consumption level and j to denote non-numeraire alternatives with zero consumption levels. In the Marshallian demand context and the utility maximization problem, Pinjari and Bhat (2011) develop a property to differentiate between alternatives that are chosen and alternatives that are not chosen. This property of the MDCEV model extends to the expenditure minimization problem and the following relationship holds:

Property 1: *The price-normalized baseline utility of a chosen alternative is always greater than that of a alternative that is not chosen. The inverse of the Lagrangian multiplier from the expenditure minimization problem can be used to differentiate between chosen and unchosen alternatives.*

$$\frac{\psi_i}{p_i} > \frac{1}{\lambda^E} > \frac{\psi_j}{p_j} \quad \text{if } i \text{ is a chosen alternative and } j \text{ is not a chosen alternative.}$$

Proof: Using the i and j notation for chosen and not chosen alternatives, I rearrange Equations (2.F-2), (2.F-3), and (2.F-4) to yield:

$$\frac{\psi_1}{p_1}(x_1)^{\alpha_1-1} = \frac{1}{\lambda^E}, \tag{2.F-8}$$

$$\frac{\psi_i}{p_i}\left(\frac{x_i}{\gamma_i}+1\right)^{\alpha_k-1} = \frac{1}{\lambda^E}, \text{ if } x_i > 0, \text{ and} \tag{2.F-9}$$

$$\frac{\psi_j}{p_j} < \frac{1}{\lambda^E}, \text{ if } x_j = 0. \tag{2.F-10}$$

Setting Equations (2.F-8) and (2.F-9) equal to each other and solving for ψ_i/p_i yields.

$$\frac{\psi_i}{p_i} = \frac{\psi_1}{p_1}(x_1)^{\alpha_1-1}\left(\frac{x_i}{\gamma_i}+1\right)^{1-\alpha_i} \tag{2.F-11}$$

The last bracketed term in Equation (2.F-11) is always greater than 1 because $x_i > 0$. I use Equations (2.F-8), (2.F-10), and (2.F-11) to write the following inequality:

$$\frac{\psi_j}{p_j} < \frac{1}{\lambda^E} = \frac{\psi_1}{p_1}(x_1)^{\alpha_1-1} < \frac{\psi_1}{p_1}(x_1)^{\alpha_1-1}\left(\frac{x_i}{\gamma_i}+1\right)^{1-\alpha_i} = \frac{\psi_i}{p_i}. \tag{2.F-12}$$

Simplifying the inequality to yields,

$$\frac{\psi_j}{p_j} < \frac{1}{\lambda^E} < \frac{\psi_i}{p_i}. \tag{2.F-13}$$

In the Marshallian demand context of Pinjari and Bhat (2011), the Lagrangian multiplier for the utility maximization problem (λ^U) is used to discriminate between alternatives that are consumed and alternatives that are not consumed.[54] In the expenditure minimization context with Hicksian demands considered in this paper, the inverse of the Lagrangian multiplier for the expenditure minimization problem λ^E is used to discriminate between alternatives. The result is consistent with duality theory which implies

[54]While the new approach uses λ^E to differentiate between consumed and unconsumed alternatives, given Equation (2.F-2), either x_1 or λ^E are sufficient for this task.

$\lambda^U = 1/\lambda^E$.

Having shown how the value of λ^E can be used to discriminate between alternatives that are consumed and alternatives that are not consumed in the expenditure minimization problem, I now derive the optimal Hicksian consumption levels and utility levels as a function of λ^E.

I first rewrite Equations (2.F-2) and (2.F-3) to yield the optimal Hicksian consumption levels for consumed alternatives as,

$$\begin{aligned} x_1 &= \left(\frac{p_1}{\lambda^E \psi_1}\right)^{\frac{1}{\alpha_1 - 1}}, \text{ and} \\ x_i &= \left[\left(\frac{p_i}{\lambda^E \psi_i}\right)^{\frac{1}{\alpha_i - 1}} - 1\right]\gamma_i, \ \text{ if } \ x_i > 0. \end{aligned} \tag{2.F-14}$$

I use the i notation instead of k to clarify that these conditions hold for alternatives with positive consumption levels. I substitute Equation (2.F-14) into the baseline utility constraint into Equation (2.F-5), which after simplification yields,

$$\bar{U} = \sum_{i=2}^{I} \frac{\gamma_i}{\alpha_i}\psi_i \left[\left(\frac{p_i}{\lambda^E \psi_i}\right)^{\frac{\alpha_i}{\alpha_i - 1}} - 1\right] + \frac{\psi_1}{\alpha_1}\left(\frac{p_1}{\lambda^E \psi_1}\right)^{\frac{\alpha_1}{\alpha_1 - 1}}. \tag{2.F-15}$$

With Equations (2.F-14) and (2.F-15) in hand, I now introduce two versions of the algorithm for calculating Hicksian demands that vary in their utility specification.

Hicksian demand algorithm with γ-profile utility functions.

The first algorithm is suitable for utility specifications that constrain all the α parameters to be equal across all choice alternatives (i.e. $\alpha_k = \alpha_1 = \alpha$). This "$\gamma$-profile" utility function (Bhat, 2008) allows us to derive analytical expressions for the Lagrangian multiplier λ^E as well as the Hicksian consumption demands for alternatives with positive consumption levels.

Equation (2.F-15) can be simplified by multiplying both sides by α and collecting like terms to yield,

$$\alpha\bar{U} + \sum_{i=2}^{I}\gamma_i\psi_i = \sum_{i=2}^{I}\gamma_i\psi_i\left(\frac{p_i}{\psi_i}\right)^{\frac{\alpha}{\alpha - 1}}\left(\frac{1}{\lambda^E}\right)^{\frac{\alpha}{\alpha - 1}} + \psi_1\left(\frac{p_1}{\psi_1}\right)^{\frac{\alpha}{\alpha - 1}}\left(\frac{1}{\lambda^E}\right)^{\frac{\alpha}{\alpha - 1}}.$$

I then solve the expression for λ^E,

$$\frac{1}{\lambda^E} = \left[\frac{\alpha \bar{U} + \sum_{i=2}^{I} \gamma_i \psi_i}{\sum_{i=2}^{I} \gamma_i \psi_i \left(\frac{p_i}{\psi_i}\right)^{\frac{\alpha}{\alpha-1}} + \psi_1 \left(\frac{p_1}{\psi_1}\right)^{\frac{\alpha}{\alpha-1}}} \right]^{\frac{\alpha-1}{\alpha}} . \tag{2.F-16}$$

I substitute the expressions for λ^E back into the rearranged first-order conditions in Equation (2.F-14) to obtain the analytic expressions for Hicksian consumption levels as a function of the baseline utility and prices:

$$x_1 = \left(\frac{p_1}{\psi_1}\right)^{\frac{1}{\alpha-1}} \left[\frac{\alpha \bar{U} + \sum_{i=2}^{I} \gamma_i \psi_i}{\sum_{i=2}^{I} \gamma_i \psi_i \left(\frac{p_i}{\psi_i}\right)^{\frac{\alpha}{\alpha-1}} + \psi_1 \left(\frac{p_1}{\psi_1}\right)^{\frac{\alpha}{\alpha-1}}} \right]^{\frac{1}{\alpha}}, \tag{2.F-17a}$$

$$x_i = \left[\left(\frac{p_i}{\psi_i}\right)^{\frac{1}{\alpha-1}} \left[\frac{\alpha \bar{U} + \sum_{i=2}^{I} \gamma_i \psi_i}{\sum_{i=2}^{I} \gamma_i \psi_i \left(\frac{p_i}{\psi_i}\right)^{\frac{\alpha}{\alpha-1}} + \psi_1 \left(\frac{p_1}{\psi_1}\right)^{\frac{\alpha}{\alpha-1}}} \right]^{\frac{1}{\alpha}} - 1 \right] \gamma_i, \text{ if } x_i > 0. \tag{2.F-17b}$$

Equations (2.F-17a) and (2.F-17b) show that once I have calculated the baseline utility and determined which subset of alternatives are consumed in positive amounts, I can derive the optimal Hicksian demands. These expressions can be then used to derive the welfare measures using Equation (2.F-7).

Specific steps in algorithm for γ-profile utility functions:

Step 0: Assume that only the numeraire alternative is chosen and let the number of chosen alternatives equal one (I=1).

Step 1: Using the data, model parameters, and either conditional or unconditional simulated error term draws, calculate the price-normalized baseline utility values (ψ_k/p_k) for all alternatives. Sort the K alternatives in the descending order of their price-normalized baseline utility values. Note that the numeraire alternative is in the first place. Go to step 2.

Step 2: Compute the value of λ^E using Equation (2.F-16). Go to step 3.

Step 3: If $\frac{1}{\lambda^E} > \frac{\psi_{I+1}}{p_{I+1}}$, go to step 4. Else if $\frac{1}{\lambda^E} < \frac{\psi_{I+1}}{p_{I+1}}$, set $I = I + 1$. If $I < K$, go back to step 2. If $I = K$, go to step 4.

Step 4: Compute the optimal Hicksian consumption levels for the first I alternatives in the above descending order using Equation (2.F-14). Set the remaining alternative consumption levels to zero and stop.

Hicksian demand algorithm with general utility functions.

In the more general case of a utility function with different α_k parameters for each alternative, I modify the algorithm in the previous section. In this context, there is no closed-form expressions for λ^E and I need to conduct a numerical bisection routine. Let $\hat{\lambda^E}$ and $\hat{U}$ be estimates of λ^E and U and let tol_λ and tol_U be the tolerance levels for estimating λ^E and U that can be arbitrarily small.

Specific steps in algorithm for general utility functions:

Step 0: Assume that only the numeraire is chosen and let the number of chosen alternatives equal one (I=1).

Step 1: Using the data, model parameters, and either conditional or unconditional simulated error term draws, calculate the price-normalized baseline utility values (ψ_k/p_k) for all alternatives. Sort the K alternatives in the descending order of their price-normalized baseline utility values. Note that the numeraire is in the first place. Go to step 2.

Step 2: Let $\frac{1}{\hat{\lambda^E}} = \frac{\psi_{I+1}}{p_{I+1}}$ and substitute $\hat{\lambda^E}$ into Equation (2.F-15) to obtain an estimate of $\hat{U}$.

Step 3: If $\hat{U} < \bar{U}$, go to step 4. Else, if $\hat{U} \geq \bar{U}$, set $\frac{1}{\lambda_l^E} = \frac{\psi_{I+1}}{p_{I+1}}$ and $\frac{1}{\lambda_u^E} = \frac{\psi_I}{p_I}$. Go to step 5.

Step 4: Set $I = I + 1$. If $I < K$, go to step 2. Else if $I = K$, set $\frac{1}{\lambda_l^E} = 0$ and $\frac{1}{\lambda_u^E} = \frac{\psi_K}{p_K}$. Go to step 5.

Step 5: Let $\hat{\lambda^E} = (\lambda_l^E + \lambda_u^E)/2$ and substitute $\hat{\lambda^E}$ into Equation (2.F-15) to obtain an estimate of $\hat{U}$. Go to step 6.

Step 6: If $|\lambda_l^E - \lambda_u^E| \leq\ tol_\lambda$ or $|\hat{U} - \bar{U}| \leq\ tol_U$, go to step 7. Else if $\hat{U} < \bar{U}$, update $\lambda_u^E = (\lambda_l^E + \lambda_u^E)/2$ and go to step 5. Else if $\hat{U} > \bar{U}$, update $\lambda_l^E = (\lambda_l^E + \lambda_u^E)/2$ and go to step 5.

Step 7: Compute the optimal Hicksian consumption levels for the first I alternatives in the above descending order using Equation (2.F-14). Set the remaining alternative consumption levels to zero and stop.

Simulation results

I conduct a set of simulation exercises to assess the computational performance of these new approaches to welfare measurement in KT models relative to the existing numerical bisection algorithm. The first simulation applies the algorithms to a real data set of recreation choices and the second set uses simulated data. All simulations compare the computational speed of the numerical bisection algorithm, the γ-profile utility function algorithm, and the general utility function algorithms using the same γ-profile utility specification. To incorporate unobserved heterogeneity, welfare estimates are calculated using the conditional approach using observed choices as detailed in von Haefen and Phaneuf (2005).[55]

Application using real data on recreation fishing choices

The first simulation exercise uses data from Lloyd-Smith et al. (2017) on recreational fishing trips in the Gulf of Mexico.[56] The choice set contains two types of fishing trips (partial and full day) in four time periods (Winter, June, Summer, Fall) for a total of 8 different types of fishing trips, as well as the numeraire (i.e. K=9). The average number of fishing trips is 3.6 while the average number of different fishing trips chosen is 1.7. The policy scenario is a \$50 price increase across all fishing trips which represents around a 5% increase on average. I vary the number of individuals (1,000 and 2,000) and error draws per individual (100 and 500).[57] All computations are coded and executed in Matlab.[58] Table 2.F.1 shows the computation times of the three algorithms to calculating welfare measures. All three approaches calculate the average annual Hicksian welfare measure for the \$50 price increase to be -\$158.44 per angler. The results show the γ-profile utility function algorithm is around 12 times faster than the numerical bisection algorithm while the more general utility function algorithm is around 3 times faster. These relative

[55] If the unconditional approach to welfare measurement is used instead, the consumption patterns in the baseline state would have to be simulated in addition to the consumption patterns under the policy state. For simplicity, I use the conditional approach which only calculates these consumption patterns once under the policy state.

[56] Additional details on the data is provided in Lloyd-Smith et al. (2017).

[57] Using the conditional approach to welfare measurement, the simulated error draws for chosen alternatives are known for certain. However, for alternatives that are not chosen, errors are drawn from truncated multivariate logistic distribution which necessitates multiple error draws per individual.

[58] All computations were performed on a desktop computer of 3.6 GHz processing speed and 8GB Random Access Memory (RAM). The speed performance of the algorithms using numerical bisection routines are sensitive to the chosen tolerance level. For this simulation, I used a tolerance level so that the welfare measures are not different at the second decimal point level.

computation times are consistent across the number of individuals and number of error draws.

As a comparison, the Pinjari and Bhat (2011) algorithm can be used to calculate Marshallian demands. The computational performance of Pinjari and Bhat (2011) algorithm is very similar (e.g. under 5% difference) to the time it takes to calculate Hicksian demands using the approach developed in this paper because of the similar analytical steps. Furthermore, the Marshallian demands can be used to calculate a welfare measure for comparison with the Hicksian measure. The 'Rule of Half' (RoH) approach to approximate consumer surplus is used (Laird, 2010), which yields an average welfare measure of -\$159.67, a difference of less than 1% from the Hicksian measure.[59] We would expect the difference to be small in this context because the price change is small, and thus the RoH approach should yield a closer estimate of consumer surplus, and income effects are small as consumer surplus is a small percentage of income, which is almost \$150,000 in the data.

Table 2.F.1: Computational performance of three algorithms to calculating welfare measures

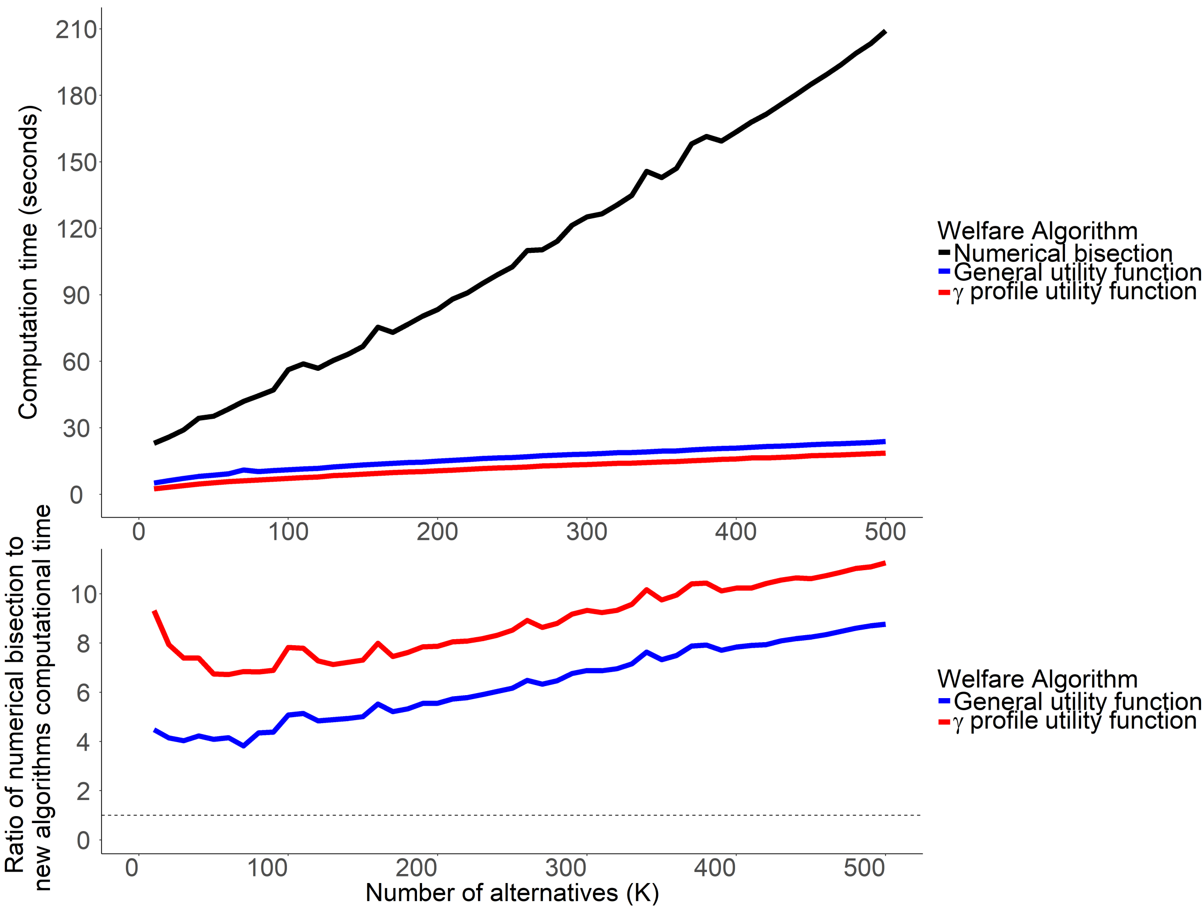
		Existing approach	New approaches	
		Numerical bisection	γ-profile utility function	General utility function
			time in seconds	
Sample size:	1000 individuals			
Number of	100 sets	49.2	3.9	14.7
error draws	500 sets	250.2	19.1	73.2
Sample size:	2000 individuals			
Number of	100 sets	98.2	7.7	29.3
error draws	500 sets	499.4	38.4	147.5

Application using simulated data

The recreation fishing data set used in the previous simulation is characterized by a relatively small number of choice alternatives (K = 9) as well as a small number of different chosen alternatives (I = 2.7 on average). I conduct two additional simulation studies that vary K and I to assess the robustness of the different algorithms under these varying choice settings. The two simulation studies use the same utility structure. The

[59] The RoH approach is defined as $RoH = 1/2\sum_k(x_k^{m0} + x_k^{m1})(p_k^0 - p_k^1)$ where x_k^{m0} and x_k^{m1} are the Marshallian demands for alternative k.

Figure 2.F.1: Simulations comparing computation performance of welfare algorithms with varying number of alternatives



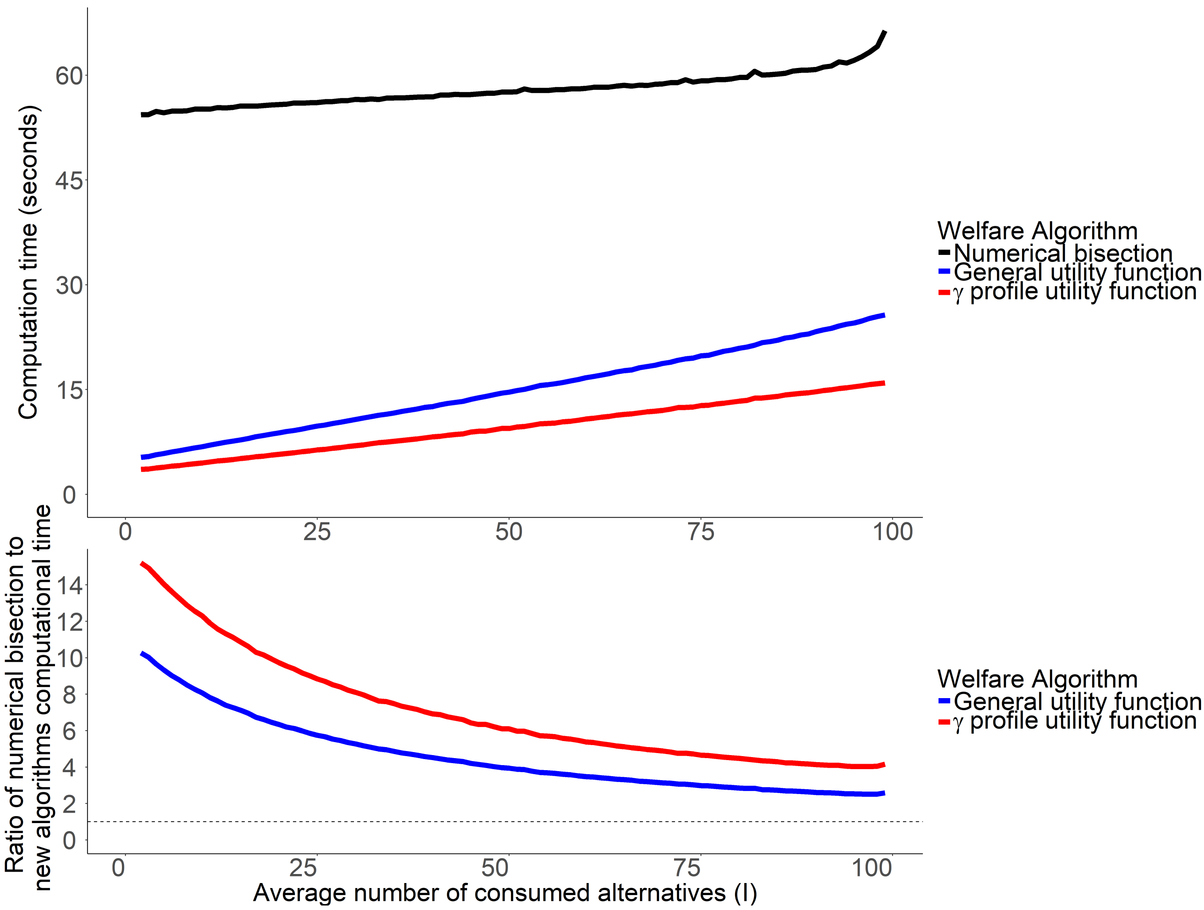
Notes: The numerical bisection routine is the existing algorithm and the general utility function and γ-profile utility function algorithms are the new approaches. The dashed line in the lower panel represents equivalent computation performance of the new approaches with the existing numerical bisection routine.

ψ term is characterized as ψ =exp$(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon)$ where x_1, x_2, and x_3 are explanatory variables affecting the baseline utility of the non-numeraire alternatives[60], the β's are parameters ($\beta_0 = -5, \beta_1 = 1.0, \beta_2 = 1.5, \beta_3 = 2.0$), and ε is the error term. The γ parameters for all non-numeraire alternatives are set to 50 and the α parameters for all alternatives are set equal to 0.5. I assume a consumption budget of $y = 100,000$ and use the utility structure to generate baseline consumption data for 1,000 individuals. The generated baseline consumption data and parameter estimates are used to analyze the welfare impacts of a policy scenario using the three different algorithms. The specific policy scenario is a 5% increase in the price of all non-numeraire alternatives. Welfare

[60]The data for these variables are generated assuming a uniformly distributed interval $[0, 2]$.

estimates are calculated using the conditional approach and assuming 50 error draws per individual.

Figure 2.F.2: Simulations comparing computation performance of welfare algorithms with varying number of chosen alternatives



Notes: The numerical bisection routine is the existing algorithm and the general utility function and γ-profile utility function algorithms are the new approaches. The dashed line in the lower panel represents equivalent computation performance of the new approaches with the existing numerical bisection routine.

For the first simulation study, the number of alternatives (K) is set to range form 10 to 500, increasing at increments of 10 for a total of 50 different simulations per algorithm. The prices of the non-numeraire alternatives are set to 100.[61] Figure 2.F.1 presents the results of the first simulation study. The top panel shows the computation time for the three algorithms. The existing numerical bisection routine computation time increases as the number of alternatives increases, while the two new algorithms introduced in this paper show more muted increases in time. To more clearly assess the relative performance of the algorithms, the bottom panel presents the ratio of the computation time of the

[61]The price of the numeraire is set to one for both simulation studies.

numerical bisection routine algorithm to the new algorithms. This ratio ranges from 3.8 to 8.8 for the general utility function algorithm and from 6.7 to 11.3 for the γ-profile utility function algorithm. For both of the new algorithms, the computational time performance relative to the numerical bisection algorithm improves once the choice set increases beyond 50.

For the second simulation study, the number of alternatives is set to 100 and the average number of chosen alternatives (I) in the data set ranges from 2 to 100.[62] Figure 2.F.2 presents the same two panels as Figure 2.F.1 for the second simulations study. The top panel shows that the computational time of the numerical bisection routine is relatively stable as I increases while the time increases linearly for the new approaches. Consequently, the computational time of the numerical bisection algorithm relative to the new approaches decreases as I increases as shown in the bottom panel. For the general utility function algorithm the ratio decreases from 10.2 when $I = 2$, to 2.5 when $I = 100$. In the case of the γ-profile utility function algorithm the ratio decreases from 15.2 when $I = 2$ to 4.1 when $I = 100$.

Conclusion

In this paper, I introduce a new approach to calculating welfare measures in KT demand models. The existing approach using a numerical bisection routine is iterative in nature and can place a large computational burdens that in some applications has significant costs. Simulations conducted in this paper using real and simulated data sets show the relative computational performance of two versions of the new approach under a range of choice contexts. The new approach reduces computational time substantially in all choice contexts investigated and performs best relative to the numerical bisection routine if i) a γ-profile utility function is specified, ii) the number of choice alternatives available is high, or iii) the average number of chosen alternatives is low. A limitation of the new approach, shared by the numerical bisection algorithm, is that it relies on an additively separable utility specification which limits the types of substitution between alternatives.[63]

Examples of when the new approach has been usefully applied include Lloyd-Smith et al. (2017) who study recreation demand. This paper required the calculation of 4.88 x

[62]To vary the number of chosen alternatives while keeping the total number of alternatives available fixed, I vary the price of the non-numeraire alternatives from 1 to 800.

[63]Bhat and Pinjari (2014) identify the use of more flexible utility structures as an important future direction for KT models and cite work by Vasquez Lavin and Hanemann (2008) as one possible approach to accommodate this.

10^8 welfare measures and took 9.9 hours to calculate using the γ-profile algorithm compared to 125.1 hours using the existing approach. Researchers facing similar challenges can benefit from these new algorithms which should facilitate the increased use of KT models in applied welfare analysis. The Matlab code for the MDCEV model and the various demand and welfare algorithms are available from the author's Github web page *https://github.com/plloydsmith*.

Chapter 3

Can Stated Measures of Willingness-to-Accept be Valid? Evidence from Laboratory Experiments

Willingness-to-accept (WTA) questions have been largely abandoned in stated preference (SP) empirical work in favour of eliciting willingness-to-pay (WTP) responses, mainly due to perceived unreliability of questions that ask respondents for compensation amounts. This paper reassesses whether stated WTA welfare measures can be valid in public and private good contexts. We conduct two laboratory experiments, for public and private goods, to analyze whether elicitation format, survey design and framing, and follow-up questions can generate truthful responses. For the public good experiment, we adapt the existing WTP incentive compatibility theoretical framework to the WTA context and test the theory using an experiment involving voting to accept payment to give up a public good. Results are consistent with the WTP literature and suggest that WTA values can be valid as long as responses have consequences for respondents. For the private good experiment, we compare different types of strategic behaviour on choice behaviour in a WTA setting focusing on whether respondents are motivated to affect the price or the provision of the good. The experiment includes treatments that place participants in a clear price or provision context and compares these settings with a consequential and hypothetical SP setting without an explicit price/provision framing. We find that strategic behaviour is present in SP questions eliciting values for private goods and in the direction expected by theory. Furthermore, survey framing and the

use of follow-up questions can provide bounds on the value estimates. These findings raise potential concerns with the use of non-incentive compatible elicitation mechanisms in WTA contexts. The results suggest that special consideration needs to be given to strategic behaviour in WTA surveys that value private goods, such as payments for ecosystem service (PES) programs.

3.1 Introduction

Asking people questions about their preferences has a long, controversial history in economics. There is an analogous parallel between the traditional predisposition amongst economists towards working with revealed preference (RP) data over stated preference (SP) data and the penchant among SP practitioners to estimate willingness-to-pay (WTP) rather than willingness-to-accept (WTA) values. The preferences for WTP measures by SP practitioners is evident in both stated and revealed sources. An influential source of the SP WTP elicitation format is the 1993 report of the NOAA Blue Ribbon Panel on Contingent Valuation (Arrow et al., 1993). This report outlines a set of best practice guidelines for value elicitation surveys that are "met by the best CV surveys and *need to be present* in order to assure reliability and usefulness of the information that is obtained [emphasis added]". In terms of choosing between WTP and WTA, the report states: "[t]he willingness to pay format should be used instead of the compensation required because the former is the conservative choice". Many authors have used these stated guidelines as justification for completely sidestepping WTA and instead estimating WTP in its place as evident in the relative empirical prevalence of the two measures.

The substantial empirical evidence suggests a clear 'revealed preference' for WTP studies among practitioners. A search of the 3,643 primary valuation studies in the Environmental Valuation Reference Inventory (EVRI) database finds that WTP studies outnumber WTA studies 14 to 1.[1] These differences are stark and suggest a large 'prevalence gap'. Moreover, WTA as a welfare measure is rarely challenged on purely theoretical grounds. In fact, WTP and WTA are both introduced in economics textbooks as useful theoretical welfare measures that economists use to think about and assess value (Freeman III et al., 2014).

This preference for WTP would appear innocuous if practitioners are only interested in situations where WTP is the proper welfare measure to use or if differences between

[1] The search of the EVRI website (www.evri.ca, accessed on November 30, 2016) returns 2,589 studies that provide WTP estimates versus only 181 studies that estimate WTA values.

WTP and WTA values estimates are relatively small. However, there are many contexts where eliciting WTA is clearly the conceptually more appropriate measure such as compensatory natural resource damage assessments or payment for ecosystem services (PES) schemes.[2] Furthermore, there is a large theoretical and empirical literature demonstrating the presence of a WTP-WTA gap. Theoretical explanations of the WTP-WTA gap are based on neoclassical reasons such as the income effect (Willig, 1976b), substitution effects (Hanemann, 1991), commitment costs (Corrigan et al., 2007), and behavioural economics explanations such as loss aversion (Kahneman and Tversky, 1979) and reference dependence (Knetsch, 2010).[3] The empirical evidence on the WTP-WTA gap has been summarized by a recent meta-analysis of 76 studies by Tuncel and Hammitt (2014). They find that the average ratio of WTA to WTP value estimates is around 3 across all types of goods and over 6 for environmental goods. In sum, there are compelling theoretical explanations and a wide body of empirical evidence suggesting that WTP and WTA values are not always close to each other; thus using WTP estimates as proxies for WTA values may yield misleading welfare calculations and policy advice (Interis, 2014).

The main explanation for why, even in the face of evidence suggesting a WTP-WTA gap, practitioners continue to elicit WTP in lieu of WTA responses is that the latter are perceived as unreliable.[4] The reasons for the perceived unreliability of WTA responses include the difficulties with scenario rejection and protest bids, the relative lack of respondent's experience with receiving compensation compared to making purchases, the fact that WTA responses are not bounded by income, and the strong perception of hypothetical bias in many empirical applications of WTA (Villanueva et al., 2017). Some of these reasons are unique to the WTA context but many are shared with WTP measures, although there is a worry that these factors are accentuated in a WTA context.

These reasons can be usefully grouped into two categories: the lack of incentive compatibility and non-conforming or invalid responses.[5] Incentive compatibility implies that

[2]These situations depend on the existing distribution of legal entitlements and reference points (Knetsch, 2007).

[3]It should be noted that there is also a literature suggesting the WTP-WTA gap is an experimental artifact (Plott and Zeiler, 2005). However, the WTP-WTA gap has been found to hold in market contexts outside of controlled experiments as well.

[4] Another potential explanation described in Interis (2014) is that WTP values tend to be a more 'conservative' measure of value where conservative is expressed as "the direction that is opposite to the survey sponsor's apparent interests (Mitchell, 2002)". As pointed out by Interis (2014), conservative does not necessarily imply smaller as there is no clear direction of the survey sponsors interest in many valuation studies and this argument in favour of WTP is inapplicable.

[5]The concept of incentive compatibility is grounded in the utility maximization hypothesis as is assumed by most analyses of SP data, but respondents may not respond to questions using this framework. Practitioners have identified many reasons for these nonconforming or invalid responses including 'yea' and 'nay' saying, scenario or payment vehicle rejection, attribute or attribute level nonattendance, and

it is in the respondent's best interest to reveal true preferences when making choices or responding to survey questions. In general, the elicitation of WTA is not thought to be incentive compatible.[6] For example, returning to the NOAA report, it states that "respondents would give unrealistically high answers to [willingness to accept] questions" and suggests that WTP responses are more incentive compatible compared to WTA responses. In their influential practitioner's guide to the econometrics of non-market valuation, Haab and McConnell (2002) state that SP approaches are generally perceived to be not able to elicit WTA responses because responses are not incentive compatible.

Carson and Groves (2007) describe how the incentive compatibility of the SBC question depends on whether the good is public or private. While the SBC question is incentive compatible for public goods under certain conditions, somewhat ironically, the incentive compatibility of SBC questions does not hold for private goods. In a WTP context, when facing a take-it-or-leave-it offer, the optimal response differs depending on whether the private good is an existing good on the market or a new good entering the market. Specifically, in the context for existing goods, respondents have the incentive to respond no to take-it-or-leave-it offers to indicate more price sensitivity. We can call this effect the price bias. On the other hand, respondents should say yes and indicate less price sensitivity for new goods to increase the chance of having the option to purchase the new good in the future. This effect can be called the provision bias. In most private good stated preference applications it is not clear how respondents think of these 'price' and 'provision' biases and the extent to which these two opposing effects on choice behaviour cancel each other out is an open question. We re-examine these issues in the WTA context by both designing treatments that place participants in price or provision contexts and by using follow-up questions to identify the participant's motivation.

---

social desirability bias. There is a general impression that these types of responses may be more prevalent in WTA contexts (Bush et al., 2013).

[6]While the majority of evidence on the incentive compatibility of different elicitation mechanisms are derived in WTP contexts, some overall empirical evidence of the incentive compatibility of WTA comes from three meta-analyses that seek to quantify and explain the reasons for the oft-cited WTP-WTA disparity. These synthesis articles define an incentive compatible mechanism as an open-ended question with a Vickrey auction, a Becker-DeGroot-Marschak (BDM) mechanism, a single binary choice (SBC) question, or a real payment mechanism. Non-incentive compatible mechanisms include simple open-ended and iterated closed-ended questions. Counterintuitively, the first meta-analysis of the WTP-WTA disparity by Horowitz and McConnell (2002) finds that studies that use incentive-compatible elicitation mechanisms have a significantly higher discrepancy. Subsequent meta-analyses by Sayman and Onculer (2005) and Tuncel and Hammitt (2014) using expanded samples and new methods both find that incentive compatible designs decrease the gap. One issue with these simple classifications of elicitation mechanisms is that, as noted above, the incentive compatibility properties depend on the specific context or good valued. For example, it is known that SBC questions are generally not incentive compatible for public goods with voluntary payments or for private goods (Carson and Groves, 2007).

The purpose of this paper is to assess whether stated WTA can be rescued as a useful empirical welfare measure. To study these issues, we examine the validity of WTA value responses elicited in both public and private good experiments. Both experiments use the participant's time as the good to be valued which mitigates "house money" effects (Harrison, 2007). In the public goods experiment, we examine the SBC format which has been shown both theoretically and empirically to be incentive compatible in the WTP context (Vossler et al., 2012; Carson et al., 2014). We show that the mechanism perspective of incentive compatibility formalized by Vossler et al. (2012) can easily be extended to the WTA context and derive analogous conditions for incentive compatibility. In many practical contexts, these conditions are less restrictive than their WTP counterparts but also raise several unique concerns. We provide the first test of the theory in a WTA context using an experiment involving voting to accept payment to give up a public good using the SBC format. Results provide support for the incentive compatibility of WTA responses to SBC questions for public goods as long as responses have consequences for respondents.

For the private good experiment, we assess and compare different types of strategic behaviour on decisions in a WTA context, focusing on whether respondents are motivated to affect the price or the provision of the good. The overall spirit of the private good experiment is to mimic a SP survey for setting up a PES scheme. We choose a PES scheme, as this is a context where WTA is the only meaningful measure: it is clearly a private good (i.e. landowner has a voluntary choice to engage in a contract involving payment for private action)[7], and there are several examples in the literature of such elicitation (Porras and Hope, 2005; Horne, 2006; Espinosa-Goded et al., 2010; Southgate et al., 2010; Kaczan et al., 2013). In the experiment, the program is the opportunity to receive payment to give up an hour of time at a later date. We use this program as it is a good that all individuals "own", contracts can be written to implement the transaction, and we can determine to offer the contract or not to respondents. Participants are first presented with a SP question and are then given the opportunity to submit an offer to participate in the program.

There are three main results for the private good experiments. First, strategic behaviour is present in SP questions for private goods for some but not all participants and follow-up questions can be helpful in identifying strategic behaviour. Second, strategic behaviour

[7]The WTA PES schemes can be considered private goods from a landowner's perspective because private actions (i.e. restoring a wetland) are being compensated with individual payments and both are excludable and rival. Of course, the benefits of restoring a wetland may have broader public good-like benefits.

biases are in the directions expected by theory with price (provision) motivated individuals less (more) likely to accept offers. Third, explicit framing of the questionnaire and the use of follow-up questions can provide useful bounds on value estimates. These findings raise potential concerns with the use of non-incentive compatible elicitation mechanisms in WTA contexts and provides an alternative approach to eliciting values in private good contexts. The results suggest that special consideration needs to be given to strategic behaviour in WTA surveys that value private goods, such as PES programs.

3.2 Experiment Implementation

We conducted two experiments at a University using participants from the University's Experimental Database. Participants are a mix of undergraduate and graduate students from all Faculties as well as campus staff. A total of 13 experimental sessions were conducted involving 202 participants.[8] The experiments were conducted between March 4 and March 16, 2016. Subjects only participate in the experiments once. The private and public good experiments were conducted as part of the same session, and for all sessions the private good experiment was conducted first, followed immediately by the public good experiment.[9]

One of the potential problems with economic experiments is the influence of money/goods provided by the experimenters. Participants may treat participation fees or goods that are given differently than their own money/goods in making decisions (Harrison, 2007). This 'house money' or 'gift' effect may limit the researcher's ability to generalize results from the laboratory setting to the outside world. There is empirical evidence suggesting individuals treat small one-time financial gains differently than their regular income (Keeler et al., 1985). Thaler and Johnson (1990) attribute these effects to a 'mental accounting' framework where small, one-time windfall gains are placed in a 'mad money' account with a higher marginal propensity to spend. Harrison (2007) finds that house money does affect behaviour of participants in a public good contribution experiment. One approach to side-stepping these issues in WTA experiments is to use goods that respondents already own (Cash, 2015). This approach can work well in private goods contexts because participants can voluntarily accept money to give up the good, but faces challenges for public goods.

[8]Experiment group sizes ranged from 12 to 19 and respondents were paid a $20 show-up fee. To minimise possible experimental biases, all experiments were conducted by the same individual, wearing the same shirt.

[9]We conducted the public good experiment second because participants had the collective choice to end the experiment early.

To mitigate these gift effects in the experiments, we use the participant's own time as the good to be given up in exchange for money. The private good considered in this experiment is to give up one hour of the participant's time at a within one week of the experiment's date. During the hour, participants would be working for one hour in the department's library helping sort and organize books. In the public good setting, we cannot coerce respondents to give up their time outside of the experiment. As a solution, we use the last 30 minutes of time in the experiment as the public good to be valued.

3.3 WTA for Public Goods

For public goods, the SBC referendum style question has long been held up as an example of an incentive compatible elicitation mechanism (Mitchell and Carson, 1989; Harrison, 2006).[10] However, it was not until Vossler et al. (2012) that a formal model was developed of the incentive compatibility properties of binary choice questions that incorporated the recent insights on the importance of consequentiality. Vossler et al. (2012) develop an explicit game theoretic model to describe the conditions informally sketched out in Carson and Groves (2007) under which respondents have the incentive to reveal their true preferences. In the WTP context, they identify four sufficiency conditions for truthful voting between a single project and the status quo: (i) the participants care about the outcome; (ii) the authority can enforce payments by voters; (iii) the elicitation involves a yes or no vote on a single project; and (iv) the probability that the proposed project is implemented is weakly monotonically increasing with the proportion of yes votes.

The first two conditions ensure that at least the costs and possibly the attributes of the project enter the respondent's utility function. The third condition eliminates the possibility for votes for one project to affect the probability of different projects being implemented. The fourth condition states that a yes vote increases the probability of the project being implemented at least some of the time. Carson et al. (2014) extend Vossler et al. (2012) theoretical results by relaxing the expected utility assumption and

[10]Besides the SBC question, there has been little empirical research into the validity of eliciting WTA measures for public goods under alternative elicitation mechanisms. Bush et al. (2013) apply concepts of mechanism design to develop a provision point mechanism (PPM) for eliciting WTA values in a contingent valuation (CV) study in Uganda. The results suggest that the PPM reduces WTA estimates by approximately 20% compared to an open-ended question format. However, as the authors note, the mechanism is not truly incentive compatible as each individual has an explicit power to veto the entire program. Messer et al. (2010) introduce the Random Price Voting Mechanism (RPVM), which is the public goods extensions of the BDM mechanism, as an incentive compatible laboratory elicitation mechanism. The RPVM operates with each participant submitting a bid or claim and if the majority of bids/claims are above a randomly drawn amount, the project proceeds.

find that the incentive compatibility properties hold.[11] Both Vossler et al. (2012) and Carson et al. (2014) provide empirical evidence using field experiments supporting the theoretical results on the incentive compatibility of WTP.

Neither Vossler et al. (2012) nor Carson et al. (2014) discuss WTA contexts, but we can extend the theoretical model by changing condition (ii). In the WTA context, the sufficient conditions for a truthful vote according to the participant's preference between a single project and the status quo are:

(i) the participants care about the outcome;

(ii) the authority can enforce voters to give up the good;

(iii) the elicitation involves a yes or no vote on a single project; and

(iv) the probability that the proposed project is implemented is weakly monotonically increasing with the proportion of yes votes.

Condition (ii) in the WTA context is the authorities can enforce voters to give up the good rather than payment as in the WTP context. Theoretically, this condition ensures that the outcomes are binding for participants and the good (or its attributes) and payment enter the participant's utility function. Practically, however, there can be important differences in the restrictiveness of this condition. In the WTP case, condition (ii) is actually quite restrictive as different payment vehicles may not be binding for different people. For example, if the payment vehicle is income tax increases, many individuals do not pay income taxes and thus may not view these costs as real for themselves. Thus, payment consequentiality has also been raised as an important component of the incentive compatibility of survey responses (Herriges et al., 2010). On the surface, condition (ii) in the WTA also appears quite restrictive as the authority may not have the ability to take certain goods away from people against their will. However, many environmental goods are public goods such as air quality or natural areas and the authorities already effective own or control the access to the good or the quality of the good. In other cases where the authorities have less control over the good, legally enforceable contracts can be written for payment. Appendix 3.A formally extends the incentive compatibility of the SBC to the WTA context.

[11] Carson et al. (2014) also prove that moving from a binding referendum to an advisory referendum or survey does not change the incentive compatibility of the mechanism.

3.3.1 Public Good Experiment Structure

The format of the public good experiment adapts the WTP experiment described in Carson et al. (2014) to the WTA context.[12] We follow the script in Carson et al. (2014) as closely as possible to provide comparable results. The good in our experiment is the participant's time and participants are told they would collectively decide what to do with the last 30 minutes of time in the experiment.[13] A vote is taken on whether everyone in the group would spend 30 minutes filling out a survey or leave immediately after the vote.

The experiment consists of three treatments.[14] In the baseline "real" treatment, participants are told that if more than 50% of the people in the group voted in favour, then everyone would receive $3 to give up 30 minutes of time to fill out the survey. If 50% or fewer people in the group voted no, then everyone could leave immediately after the vote and receive no additional money. In the hypothetical treatment, the vote is not binding and participants are told that "*regardless of the vote outcome, no one will receive an additional $3 or have to stay the extra 30 minutes to fill out the questionnaire*". The "consequential" treatment included a probabilistic referendum that set the probability that the referendum would be binding to 50%. In this treatment, a two-step referendum format is used. The first step consists of participants voting using the same simple majority voting rule to determine if the referendum passes as the real treatment. If the referendum passes, the second step determines if the referendum is binding with a flip of a coin. If the referendum binds, every participant is paid the $3 and has to give up the 30 minutes of time to fill out the survey. If the referendum does not pass or does not bind, no compensation is paid and all participants can leave immediately after the vote.

3.3.2 Hypotheses and Analysis

The experimental design allows us to test several hypotheses pertaining to the response of voting behaviour to the different treatments. The three hypotheses to be tested are formally stated and then discussed. Let p be the probability that the treatment is consequential.

[12] The full public good experiment instructions are provided in Supplementary Material C.

[13] Participants were not provided with any details on the content of the survey. In experimental sessions where the majority voted yes and the referendum was binding, all of the 46 dissenting participants who voted no to giving up their time stayed for the duration of the experiment which indicates that the respondents perceived the votes to be consequential.

[14] We conducted four sessions of the real and consequential treatment, and five sessions of the hypothetical treatment.

1. Hypothesis 1: The percentage in favour at p = 0.5 (consequential) is equal to that of p = 0 (hypothetical)

2. Hypothesis 2: The percentage in favour at p = 0.5 (consequential) is equal to that of p = 1 (real)

3. Hypothesis 3: The percentage in favour at p >0 (consequential and real) is equal to that of p = 0 (hypothetical)

In light of the incentive compatibility framework, we expect to reject Hypothesis 1 as hypothetical (i.e. inconsequential) and consequential voting behaviour need not be the same. We expect to not reject Hypothesis 2. According to the theory of consequentiality, participants should treat consequential and real votes similarly. We expect to reject Hypothesis 3 for the same reasons as Hypothesis 1.

3.3.3 Public Good Experiment Results

Table 3.1 presents the voting distribution for the three treatments. The voting results show that only 46.8% (37/79) voted yes in the hypothetical treatment compared to 63.9% (39/61) in the consequential treatment and 69.4% (43/62) in the real treatment. The near 50-50 voting split for the hypothetical treatment suggests that participants may have behaved as though they are flipping a coin due to the inconsequential nature of the decision. This inconsequential voting pattern has been found in previous studies (Cummings et al., 1997; Burton et al., 2007).

Table 3.1: Public good experiment voting summary statistics

Treatment	Number of Participants	Vote (%)
Hypothetical	79	46.8
Consequential	61	63.9
Real	62	69.4

To statistically test our hypotheses, we use different subsets of the data to estimate probit regression models of the probability of a yes response against different treatment dummy variables. Table 3.2 shows the results for the three hypotheses. The first column compares the hypothetical and consequential treatment data and includes a dummy variable for the hypothetical treatment to test Hypothesis 1. We reject the hypothesis at the 0.05 level that responses in the two treatments are equal. The second column in Table 3.2 shows the results for Hypothesis 2. The results show that there are no statistically significant differences in voting behaviour between the consequential and real treatments. The final column of Table 3.2 includes data for all three treatments and a single dummy variable for the responses from the hypothetical treatment. Results are consistent with Hypothesis 3 and show that voting behaviour in the hypothetical treatment is statistically significant different from voting behaviour in the other two treatments.

Table 3.2: Probit regression results for public good experiment hypotheses

	Hypothesis 1	Hypothesis 2	Hypothesis 3
Hypothetical treatment	-0.436^{**}		-0.510^{***}
	(0.217)		(0.183)
Consequential treatment		0.149	
		(0.234)	
Constant	0.357^{**}	0.357^{**}	0.431^{***}
	(0.164)	(0.164)	(0.117)
Treatment included			
Hypothetical	Y	N	Y
Consequential	Y	Y	Y
Real	N	Y	Y
Observations	140	123	202
Log Likelihood	−94.48	−78.09	−132.89

Note: *p<0.1; **p<0.05; ***p<0.01

Taken together, the results are consistent with our theoretical expectations regarding the importance of the incentive compatibility of WTA elicitation mechanisms. However, the results so far do not control for participant characteristics. The random assignment of respondents to treatment should render such controls unnecessary and Table 3.B.1 in Appendix 3.B reports the summary statistics of participant characteristics and pair-wise comparisons between the different treatment sub samples. As a robustness check, we estimate models with such controls and these results are presented in Table 3.B.2 in Appendix 3.B. None of the results substantially change with respondent characteristic controls.

Another consideration in comparing hypothetical and binding referendums is the variance associated with the choices (Haab et al., 1999; Carson et al., 2014). We use various probit model specifications, including the heteroskedastic probit model, to statistically test the effect of allowing the different treatments to affect both the mean and variance of the underlying WTA distributions. The results of these tests are discussed in Appendix 3.B and presented in Table 3.B.3 and corroborate the results presented in Table 3.2.

3.4 WTA for Private Goods

As mentioned above, the incentive compatibility of SBC questions does not hold for private goods in the majority of cases (Carson and Groves, 2007).[15] When facing a take-it-or-leave-it offer for a private good, respondents have the incentive to respond no to indicate more price sensitivity for existing goods (i.e. the price bias) and have the incentives to say yes to increase the chance of having the option to purchase the new good in the future (i.e. the provision bias).[16] These biases were studied in a WTP laboratory experiment by Lusk et al. (2007) who find support for strategic behaviour responses for some people.

For private goods, while Carson and Groves (2007) do not discuss incentive compatibility in the WTA context, we hypothesise that the strategic behaviour biases would be similar to their WTP counterparts. Specifically, it would be optimal to say no to questions involving existing offers for goods to increase perceived necessary compensation levels and to say yes to potential new offers for goods to increase the chance of the offers being introduced. Using a PES scheme as an example, respondents to a survey for a new scheme may have the incentive to understate their WTA to increase the likelihood of the program being implemented, for which later on they can decide whether to participate. Alternatively, if the PES scheme is already in operation and a SP survey is administered to assess the potential to tinker with the compensation levels, respondents have the incentive to overstate their WTA.

There are three types of empirical strategies available to improve the incentive compatibility of WTA responses in private goods contexts: different elicitation mechanisms (Lusk and Shogren, 2008), auction cheap talk scripts (Krishna et al., 2013; Kanjilal, 2015), and learning rounds and rationality spillovers (Chilton et al., 2011). Lusk and Shogren (2008) review the experimental literature considering the use of different elicitation mechanisms such as Vickrey auctions, BDM mechanisms and the random incentive system. There is also a substantial body of theoretical and empirical work in the field using conservation

[15]Carson and Groves (2007) note a special case where the incentive compatibility properties of SBC questions can be restored if the choice is between two different private goods and only potential users participate.

[16]The distinction between newly introduced goods and goods already in production is clearer for market goods, but the concepts remain somewhat vague for non-market goods.

auctions (Latacz-Lohmann and Van der Hamsvoort, 1997; Latacz-Lohmann and Schiizzi, 2005). Most of this work has focused on the optimal design of auction mechanisms with the aim of reducing information rent and/or improving environmental outcomes. Cheap talk scripts have also been used to increase the validity of WTA responses. Krishna et al. (2013) introduce the concept of competitive bidding into their survey of compensation payments to Indian farmers to make the responses more incentive compatible. Respondents were told that the government had a limited budget and only the least-cost providers would be selected to participate in the PES scheme.[17] Another example of the use of cheap talk script is in Kanjilal (2015) who use a description of the auction mechanism that would be used in a real program before eliciting WTA responses from farmers in Saskatchewan and Alberta.[18] Similar to the use of cheap talk scripts to reduce hypothetical bias, we do not know if incentive compatibility cheap talk scripts can completely induce truthfully responses. With the exception of auction cheap talk scripts, the main limitation with these existing approaches is that they do not easily transfer to the SP survey setting.

3.4.1 Private Good Experiment Structure

The objective of the second experiment is to assess and compare different types of strategic behaviour on choice behaviour in a WTA private good setting. The overall spirit of the private good experiment is to mimic a SP survey for setting up a PES scheme. The scenario is modeled on the introduction of a new voluntary program that will pay participants compensation for undertaking some costly activity. Participants are told how their responses to the SP survey will be used to determine whether the program is offered to the group and the expected program payment levels. If the program is offered to the group, participants then have a second choice of whether to participate in the program and at what price.

To identify strategic behaviour in WTP settings, follow-up questions have been proposed as one method to identify strategic behaviour (Lusk et al., 2007; Doyon and Bergeron, 2016). However, using these responses in statistical models of behaviour is challenging

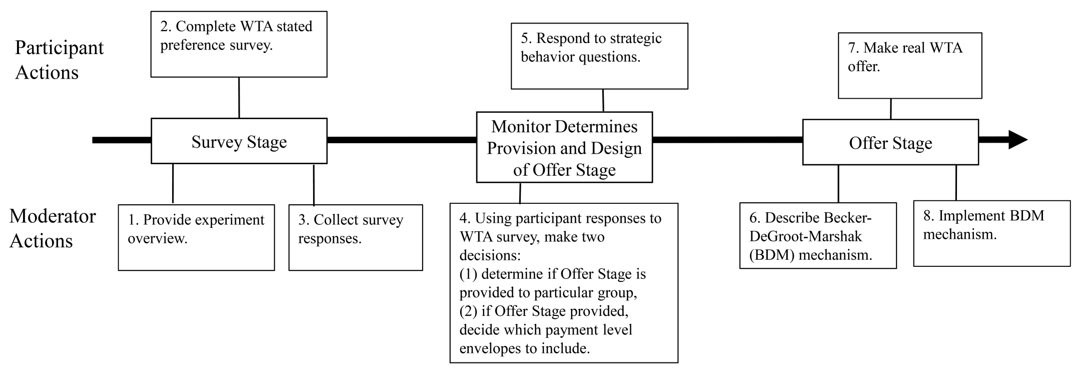
[17] The specific script used before the valuation question is: *You only need to decide whether you want to participate or not. However, before you answer, please note that only a limited number of households in the Kolli Hills would be selected to participate in this scheme, as the amount of funding for the scheme would be limited. Therefore the smaller the amount of support you would require to participate in the programme, the higher are your chances of being selected.*

[18] The specific script used is: *Suppose you were asked to submit a sealed bid that represents the amount that you will be willing to accept in compensation for participating in the program. The agency would select the winning bids by choosing the lowest bids according to their budget.*

due to potential endogeneity concerns. To complement the use of follow-up questions, we also use different treatments that introduce a clear price or provision framing to the use of the SP survey. These exogenous treatments help provide an alternative means of assessing the effects of strategic behaviour on choice decisions.

The experiment itself consists of two stages: a Survey Stage and an Offer Stage. The Survey Stage includes a SP question and a description of how responses to this question will be used by the moderators. After the Survey Stage, the moderators decide whether to provide the Offer Stage to the particular experiment group or not and what payment levels to include in the Offer Stage. If the Offer Stage is provided to the group, respondents make an offer of the minimum amount that they would be willing to accept to sell an hour of their time. The private good considered in this experiment is to give up one hour of the participant's time to work in a library sorting and organizing books at a later date. Figure 3.1 provides an overview of the two experiment stages as well as the sequential actions by participants and the moderator. We describe each stage and action in detail below. The full private good experiment instructions is provided in Supplementary Material D.

Figure 3.1: Overview of private good experiment and actions by participants and moderators



3.4.1.1 Survey Stage

The Survey Stage starts with the moderator providing a brief overview of the experiment including an overview of Offer Stage (Step 1). Participants are informed that the Offer Stage may not be provided to every group and different payment level may be presented to different groups. Participants are then asked to respond to a SP question and are told how their responses will be used (Step 2). Specifically, participants are told that

*Responses in the **Survey Stage** will help the researchers decide whether or not to present the **Offer Stage** to this group as well as determine which payment level envelopes to provide in the **Offer Stage**, depending on our budget.*

The choice question uses the stochastic payment card (SPC) approach. The SPC approach is closely related to the multiple-bound discrete choice (MBDC) but allows respondents to use a combination of words and numerical values to more easily express themselves (Wang and Whittington, 2005). We employ a SPC approach as it is well suited to the context, parallels the consequential ("real") treatment, and provides considerable value information for each individual. The SPC question used in the experiment is provided in Figure 3.2. All responses in the Survey Stage are collected by the moderator (Step 3).

3.4.1.2 Moderator Determines Provision and Design of Offer Stage

The moderator analyzes the responses to the Survey Stage and makes two decisions (Step 4). The first decision is whether to provide the Offer Stage to this particular experimental group or not. The second decision is what payment levels to include in the Offer Stage. The moderator has 15 different envelopes containing different payment amounts ranging from \$5 to \$20 and decides on 10 to include in the Offer Stage. Note that there is no explicit link between the participant's individual response to the Survey Stage and the Offer Stage. Instead, the aggregate information from the Survey Stage is used to inform the Offer Stage design.

While the moderator is analyzing the Survey Stage responses, participants are asked two different questions to identify the presence of strategic behaviour in the Survey Stage (Step 5). The first question format focuses on the motivation of the respondents:

Which of the following motives were important when deciding to accept or not accept the offer at each payment level? Select all that apply.

- *A. Ensure the offer is worth it for myself given the payment levels presented*
- *B. Ensure the Offer Stage is provided to my group*
- *C. Ensure the payment levels in the Offer Stage are favourable for myself*

If participants selected more than one motivation, they are asked a follow-up question on which one is the most important. The first motivation option identifies participants who

Figure 3.2: Example of the stochastic payment card design

Suppose you have been given the opportunity to be paid money to give up one hour of your time to work in the department's library helping sort and organize library books. The work can be completed any weekday between 9 am and 5 pm and the work must be completed within two weeks of today's date (March 24, 2016). Please think about a suitable time when you would complete the work before you respond to the question. You would be paid once you've completed the hour of time.

How likely would you be to accept the offer to give up one hour of your time if the payment amount is....? Please select a response for each payment amount (one response per row).

Payment for 1 Hour of Time	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
\$5	☐	☐	☐	☐	☐
\$10	☐	☐	☐	☐	☐
\$15	☐	☐	☐	☐	☐
\$20	☐	☐	☐	☐	☐
\$25	☐	☐	☐	☐	☐
\$30	☐	☐	☐	☐	☐
\$35	☐	☐	☐	☐	☐
\$40	☐	☐	☐	☐	☐

are motivated by whether the program is beneficial to them, and not necessarily strategic behaviour. The second option identifies 'provision motivation' and participants who are focused on the provision of the Offer Stage. The third option identifies 'price motivation' and participants who aim to affect the payment levels in the Offer Stage.

The second strategic behaviour question format focuses on the perceived price and provision consequences of the Survey Stage choices. Using a 1 to 5 scale, where 1 is not at all and 5 is definitely taken into account, participants are asked a price consequence question, "*to what extent do you think your choices will be taken into account for determining which payment levels are provided in the Offer Stage*", and a provision consequence question, "*to what extent do you think your choices will be taken into account for determining the chances of being provided the Offer Stage*".

3.4.1.3 Offer Stage

The moderator describes the open-ended BDM mechanism with a random payment amount as the elicitation mechanism (Step 6). Participants make a monetary offer of the minimum dollar amount they would be willing to accept to give up an hour of their time as described earlier (Step 7). The offers are compared to a randomly drawn payment level contained in one of the 10 envelopes held by the moderator (Step 8). No information on the range or distribution is provided to the participants. There are two possible outcomes. Offers are accepted if the amount they indicate is less than or equal to the randomly drawn payment level. In this case, participants would receive the randomly drawn payment level in dollars and have to give up an hour of their time. Offers are not accepted if the amount they indicate is more than the randomly drawn payment level. In this later case, they do not receive any money and do not have to give up an hour of time.

3.4.1.4 Treatments

1. Consequential treatment: This treatment serves as the baseline and the text before the Survey Stage describes how responses will be used to inform the decision to present the Offer Stage to the group and the design of payment levels provided. The Offer Stage is a real binding contract that may or may not be offered to participants depending on their responses to the Survey Stage.

    - Text before Survey Stage: *Responses in the Survey Stage will help the researchers decide whether or not to present the Offer Stage to this group as well as determine which payment level envelopes to provide in the Offer Stage, depending on our budget.*

2. Hypothetical treatment: This treatment is exactly the same as the consequential treatment, but the Offer Stage is hypothetical. The following text is provided at the beginning of the Offer Stage: *You will not actually have to give up the time nor will you receive the money. But we ask that you make choices as if you were actually making a real money decision on whether to accept the payments and give up an hour of your time.*

- Text before Survey Stage: Same as consequential

3. Price framing treatment: This treatment fixes the provision of the Offer Stage and the text before the Survey Stage tells participants that their responses will be used to design the payment levels (i.e. compensation amounts) offered in the Offer Stage. The Offer Stage is a real binding contract.

    - Text before Survey Stage: *Responses in the Survey Stage will help the researchers determine which payment level envelopes to provide in the Offer Stage, depending on our budget.*

4. Provision framing treatment: In this treatment, the payment level envelopes are selected before the experiment instead of deciding which payment levels to include in the Offer Stage based on Survey Stage responses. The text before the Survey Stage tells participants that their responses will be used to determine whether or not the Offer Stage will be offered. The Offer Stage, if provided, is a real binding contract.

    - Text before Survey Stage: *Responses in the Survey Stage will help the researchers decide whether or not to present the Offer Stage to this group, depending on our budget. The payment levels potentially being provided in the Offer Stage have already been determined.*

3.4.2 Private Good Experiment Results

Table 3.3 provides a summary of participants and monetary WTA offers in the private good experiment treatments. Table 3.C.1 in Appendix 3.C presents a comparison of participant characteristics across treatments. The Monetary Offer column is the mean of the values provided by participants in the Offer Stage.[19] For all the non-hypothetical treatments, a total of 6 offers were accepted and 5 individuals followed-up to receive payment and work in the department for one hour each.[20]

[19] Comparing the monetary offers in the hypothetical treatment and the other binding treatments, we can assess whether the consequentiality of questions matter in WTA private good settings. The price, provision, and consequential treatments are exactly the same in the Offer Stage and we expect these treatments to all yield similar results. Using a conventional t-test, the mean offers of \$16.5 for the hypothetical treatments and \$18.4 for the binding treatments are significantly different at the 10% level.

[20] Due to human ethics restrictions, we could not force people to show up if their offers were accepted. While this is a potential limitation on the binding nature of the offer stage, we believe the offer stage was perceived as binding due to the fact that 5 people showed up out of the 6 accepted offers.

Table 3.3: Willingness-to-accept to give up an hour of time in the private good experiment

Treatment	Number of Participants	Mean Monetary WTA Offer	Range of WTA Offers (min - max)
Consequential	65	$19.2	($1 to $40)
Hypothetical	42	$16.5	($5 to $30)
Price framing	49	$17.9	($5 to $40)
Provision framing	46	$17.8	($5 to $30)
Total	202	$18.0	($1 to $40)

To assess the issues surrounding strategic behaviour in the private good context, we present the results as answers to three questions:

1. To what extent is strategic behaviour present in SP questions eliciting values for private goods?

2. How does the strategic behaviour of respondents affect private good choice decisions?

3. Can the treatments and strategic behaviour questions provide informative bounds on WTA?

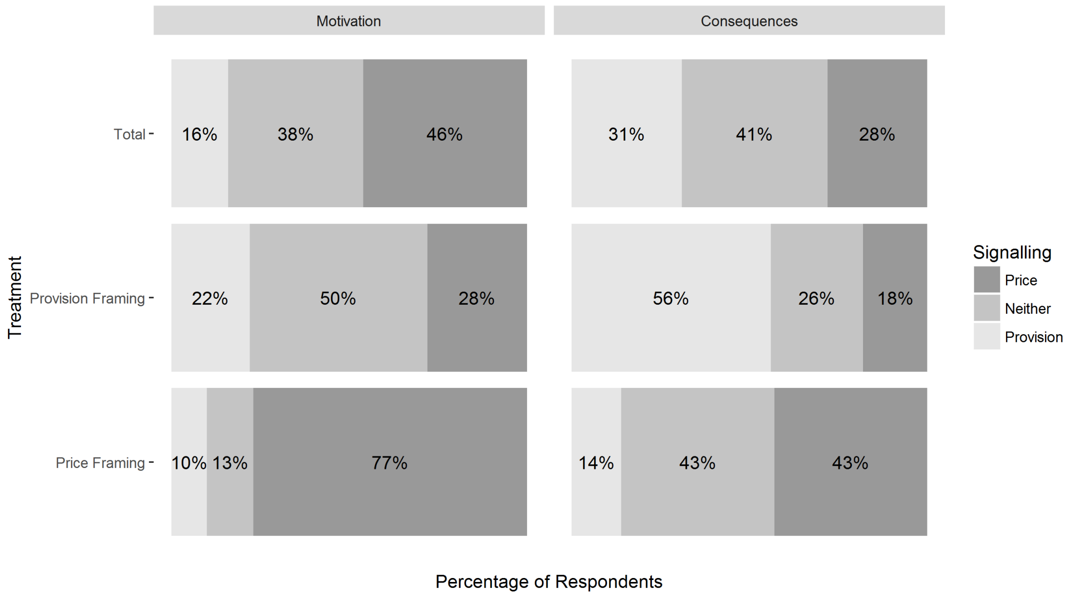
To what extent is strategic behaviour present in SP questions eliciting values for private goods?

To answer this question, we examine how the responses to the strategic behaviour follow-up questions varied by treatment.We compute the percentage of respondents who perceive their responses as influencing the prices or provision of the Offer Stage and summarize the results in Figure 3.3. The left panel of Figure 3.3 reports the percentage of respondents who are price or provision motivated in each treatment.[21] These motivation questions are intended to identify if the respondent thought their Survey Stage responses were being used primarily to ensure the Offer Stage is provided to their group (Provision motivation) or used to set the payment levels provided in the Offer Stage (Price motivation). Across all treatments, 46% of participants indicated they are motivated by price considerations while 16% indicated a provision motivation, with the remaining 38% indicating neither price nor provision motivations. As expected, the percentages of price and provision motivated individuals are highest in their corresponding treatments with 77% of respondents in the price framing treatment and 22% of respondents in the provision framing treatment indicating price and provision motivations, respectively.

[21] Respondents who indicated both provision and price motivations are recoded as zeros for both variables as we assume the competing effects net each other out.

Using the consequences follow-up questions, we code individuals as price consequential if they indicate a larger number for the price consequentiality question compared to the provision consequentiality question and vice versa for provision consequential.[22] The right panel in Figure 3 report the percentage of individuals who are price and provision consequential or neither. Around 28% and 31% of respondents are price and provision consequential, respectively, across all the treatments. As with the motivation follow-up question, the percentages of price and provision consequence focussed individuals are highest in their corresponding framing treatments.

Figure 3.3: Identifying strategic behaviour in the private good experiment



Notes: Numbers are percentages of respondents in each treatment group. The left panel uses the motivation strategic behaviour follow-up question and presents the percentage of respondents who indicated they are price or provision motivated or neither in the treatments. The right panel uses the consequences strategic behaviour follow-up question and reports the same numbers.

To more formally examine the impact of the framing treatments on responses to the two strategic behaviour follow-up questions, we estimate a set of probit models using price and provision strategic behaviour responses as the dependent variable and include dummy variables for the different treatments. Table 3.4 presents the results. For the motivation follow-up question, the provision framing treatment had a negative and significant effect on responses to the price motivation question and a positive and insignificant effect on responses to the provision motivation question. As expected, the price framing treatment had a positive and significant effect on price motivated responses and a negative and significant effect on provision motivated responses.

[22]Respondents who stated the same number for both price and provision consequence are coded as zeros for both variables.

Table 3.4: Probit models of framing treatments on strategic behavior

	Motivation		Consequences	
Parameter	Price	Provision	Price	Provision
Hypothetical treatment	-0.134	0.427*	-0.374	0.718***
	(0.25)	(0.254)	(0.276)	(0.265)
Provision framing treatment	-0.586**	0.035	-0.421	1.038***
	(0.246)	(0.254)	(0.273)	(0.260)
Price framing treatment	0.504**	-0.723**	0.322	-0.17
	(0.254)	(0.293)	(0.243)	(0.286)
Constant	0.253	-0.547***	-0.502***	-0.898***
	(0.157)	(0.164)	(0.163)	(0.181)
Observations	202	202	201	201
Log Likelihood	-129.06	-112.6	-115.09	-110.78

Note: *p<0.1; **p<0.05; ***p<0.01. For all models, the Consequential treatment is the omitted reference category.

The last two columns of Table 3.4 present the probit model results using responses to the consequences strategic behaviour follow-up questions as dependent variables. We can see that the framing treatments are not as effective in influencing the strategic behaviour of respondents compared to their motivation as captured in the motivation questions. Only the provision framing treatment has a statistically significant impact on provision consequences. Given the lack of response of the consequences follow-up question to the different framing treatments, we only present the results using the motivation follow-up questions in subsequent sections.

Taken together, these results suggest that a significant portion of the sample was behaving strategically in the Survey Stage. Furthermore, the strategic behaviour follow-up questions appear to be identifying price and provision signalling as the framing treatments affect strategic behaviour in the expected direction.

How does the strategic behaviour of respondents affect private good choice decisions?

For the second question, we investigate how the framing treatments and strategic behaviour follow-up questions affect the decision of whether to accept a given payment level in the Survey Stage. The polychotomous responses for each payment level are converted into a binary variable using the 'Probably Yes (75%)' as the lower bound cut-off for a 'yes' response. We use these converted binary responses and estimate a random effects logit model to account for multiple responses per individual.[23] Table 3.5 presents these random effects logit model results. The first column shows the results using only a constant term and the payment amount. As expected, the higher the payment amount, the more likely an individual is to give up an hour of their time. The second column includes dummy variables for the hypothetical, provision framing, and price framing treatments with the consequential treatment being the reference category. The provision framing treatment variable is positive and significant implying that participants in this treatment are more likely to accept the program. This result is consistent with the provision signaling hypothesis as respondents had an incentive to underbid in this treatment to increase the likelihood of being presented with the Offer Stage. The price framing treatment variable is close to zero and is not significant.

The third column adds the strategic behaviour follow-up questions to the model while acknowledging that these variables are potentially endogenous.[24] We also include interaction terms between these motivation responses and the payment amount because these two motivations will have different impacts on the price sensitivity of respondents. Specifically, provision motivated respondents will want to indicate less price sensitivity while price motivated respondents will want to indicate that they are more price sensitive. The results support the provision and price signaling hypotheses. The positive and significant coefficient for the provision motivation variable indicates that these respondents are more likely to accept the program while the negative and significant coefficient for the price motivation variable suggests the opposite. Furthermore, the interaction term between the payment amount and provision motivation is negative, as expected, as these respondents are less price sensitive. The positive and significant coefficient for the payment amount and price motivation interaction term suggests that these respondents are more price sensitive. Overall, it appears that strategic behaviour did influence private good choices in the directions expected by theory.

Can the treatments and strategic behaviour questions provide informative bounds on WTA?

[23] Two individuals did not respond to all payment levels in the SPC, resulting in 1,612 responses from 202 individuals. As a robustness check, we consider two alternative lower bound cut-off levels ('Not Sure (50%)' and 'Definitely Yes (100%)') and conduct the same analysis. Results are presented in Table 3.C.2 in Appendix 3.C and show a consistent qualitative pattern as the third column of Table 3.5.

[24] The use of an instrumental strategy using the exogenous treatments as instruments was investigated but the small sample size reduced the robustness of this approach.

Table 3.5: Random effects logit models of the decision to accept payment for time

Parameter	(1)	(2)	(3)
Constant	-4.919^{***}	-5.381^{***}	-5.460^{***}
	(0.370)	(0.458)	(0.634)
Amount	0.329^{***}	0.329^{***}	0.329^{***}
	(0.021)	(0.021)	(0.030)
Hypothetical treatment		0.751	0.812
		(0.460)	(0.502)
Provision framing treatment		1.203^{***}	1.265^{**}
		(0.454)	(0.492)
Price framing treatment		0.138	0.110
		(0.436)	(0.493)
Provision motivation			2.263^{***}
			(0.785)
Price motivation			-2.084^{***}
			(0.777)
Amount*Provision motivation			-0.125^{***}
			(0.038)
Amount*Price motivation			0.133^{***}
			(0.042)
Observations	1,612	1,612	1,612
Participants	202	202	202
Log Likelihood	−490.56	−486.20	−467.44
AIC	987.11	984.39	954.88

Note: *p<0.1; **p<0.05; ***p<0.01. AIC = Akaike Information Criteria. The dependent variable is whether the participant would accept the payment level using 'Probably Yes (75%)' and 'Definitely Yes (100%)' responses as yes responses. The provision and price motivation variables represent whether the participant indicated these were motivations in responding to the survey. For all models, the Consequential Treatment is the omitted reference category

For the third question, we use the different framing treatments and strategic behaviour questions to provide bounds on WTA value estimates from the Survey Stage. We consider the framing treatments without the potential endogenous follow-up questions and use the model estimates from the second column of Table 3.5. As a lower bound, the WTA for the participants in the provision framing treatment is estimated to be \$12.68 (standard error (se): 1.047). As an upper bound, the WTA for participants in the price framing treatment is estimated to be \$15.92 (se: 1.000).[25] These lower and upper bound estimates of WTA can be used to inform the potential range of WTA values.

3.5 Conclusion

Economists have largely abandoned SP WTA welfare measures in empirical work. In this paper, we reassess whether this abstinence is justified. We report the results of two WTA experiments involving private and public good contexts. Both experiments use time as the good to be valued which mitigates house money effects and provides a useful public goods context. For public goods, we adapt the WTP incentive compatibility proof outlined in Vossler et al. (2012) to the WTA contexts. These adapted incentive compatibility conditions highlight two key differences between the WTP and WTA context. First, payment consequentiality in the WTA context is quite different compared to the WTP context. Finding and defining a credible payment vehicle that applies to the population of interest has remained a challenge in the WTP context (Johnston et al., 2017). In the WTA context, these challenges may be less of an issue because people have a greater incentive to receive rather than provide payment. However, there may be remaining concerns about the government actually following through on its pledges to pay citizens or reduce their taxes. Second, the WTA context involves the government or relevant authority having the ability to enforce voters to give up the good and enforcement could be an issue. However, governments often already control the access, quality, or quantity of many public goods such as air or water quality and in many cases can credibly determine the final amount of good provided. Conversely, in the case of a damage assessment after an event, the good is effectively already taken away and the WTA question only relies on the perceived consequentiality of payment if not cleaned up. Depending on the context, these two differences can have important practical impacts on the viability of eliciting either welfare measure.

[25]We ca also use the model coefficients from the third column of Table 3.5 which includes responses to the strategic behaviour follow-up questions to construct the bounds. Participants that indicated they are provision motivated have a WTA of \$12.76 (se: 2.139) while price motivated participants have a WTA of \$15.61 (se: 0.592).

The adapted WTA incentive compatibility theory for public goods is tested in a laboratory experiment that places people in a hypothetical, consequential, or real treatment. The results provide strong support for the use of consequential SBC WTA questions as long as the probability that the responses are taken into account is greater than zero. While the experimental set-up abstracts away from many issues with eliciting WTA questions in practice, Whittington et al. (2017) provides some guidance on asking WTA questions in SP surveys.

In both the public and private good experiments, participants tended to understate their WTA in the hypothetical treatments compared to the consequential treatments. This result corroborates the study by List and Shogren (2002) which finds a similar understatement of values in their WTA experiments involving surrendering holiday gifts.

The second part of the paper focuses on eliciting WTA values for private goods. The experimental results suggest that strategic behaviour is present in SP questions eliciting values for private goods. Identifying strategic behaviour is not straightforward however and the approach in this paper used both explicit survey framing and follow-up questions. While the explicit framing approach benefits from randomly placing people in separate price and provision framing treatments, this may not be appropriate or viable in all private good contexts. The use of strategic behaviour follow-up questions suffers from many of the same issues that plague the increasing use of perceived consequentiality questions (Herriges et al., 2010). Both benefit from being relatively simple to append to a usual SP survey, but using these responses in data analyses raises issues of potential endogeneity. Furthermore, while a strategically behaving participant may not voluntarily reveal this information in the follow-up question, participants who chose to misrepresent their motivations makes detecting the influence of strategic behaviour more difficult. Conducting a large field experiment with both survey framing and follow-up questions would allow the endogeneity of these questions to be controlled for using an instrumental variable strategy.

The second main result of the private good experiment is that strategic behaviour affects private good decisions in the directions expected by theory. Participants who are price motivated are less likely to accept the program and show greater price sensitivity while provision motivated individuals are more likely to accept the program. The results also demonstrate how survey framing and follow-up questions can be used to provide bounds on WTA values.

The results of this paper have implications for the burgeoning use of SP methods in designing PES schemes. There is a lack of clarity in the literature on how PES schemes fit into the public/private good distinction and the appropriate role of SP elicitation methods. In developing PES schemes, SP methods have been used to either estimate the WTP of users of the services for the program or the WTA of landowners to accept payment.[26] Part of the confusion stems from the fact that PES schemes are often targeted at providing public goods for the users such as carbon sequestration or water quality improvements. However, from a landowner's standpoint, the payments and activities associated with these programs more closely mimic private goods.[27] As outlined above, the private/public good distinction is important for incentive compatibility and if we interpret PES schemes to be private goods for landowners then SP research has exclusively used non-incentive compatible elicitation mechanisms. For example, Liu et al. (2014) use a SP survey to elicit the WTA of landowners in Iowa for adopting perennial strips. They conduct a convergent validity test of a traditional multinomial choice question, a modified multinomial choice question suggested by Carson and Groves (2007) where all but one of the alternatives are implemented, and a SBC question.[28] For public goods, the latter two elicitation schemes are incentive compatible, but none of the three elicitation mechanisms are generally incentive compatible for private goods. Other SP research on estimating WTA for PES schemes have used elicitation mechanisms that are not necessarily incentive compatible such as binary choice questions (Southgate et al., 2010), open-ended questions (Southgate et al., 2010), and choice experiments (Porras and Hope, 2005; Horne, 2006; Espinosa-Goded et al., 2010; Kaczan et al., 2013). While there is no commonly accepted survey-based elicitation mechanism that is incentive compatible for private goods, the results of this study show how explicitly incorporating strategic behaviour framing into the design of the survey and the use of strategic behaviour follow-up questions can help inform and control for biases in value estimates.

[26] In the developing world context, Whittington and Pagiola (2012) review and assess the use of CV in the design of PES and solely focus on WTP studies in their evaluation due to a paucity of WTA studies.

[27] The WTA PES schemes can be considered private goods from a landowner's perspective because private actions (i.e. restoring a wetland) are being compensated with individual payments and both are excludable and rival. Of course, the benefits of restoring a wetland may have broader public good-like benefits.

[28] The results in Liu et al. (2014) suggest no significant difference in estimates between the two multinomial choice questions while the results diverge between the multinomial and SBC formats.

Appendix 3.A Proof incentive compatibility of single binary choice responses in WTA contexts

The WTA adaptation below heavily relies on the WTP incentive compatibility proof outlined in Vossler et al. (2012). As described in the text, the sufficiency conditions for the incentive compatibility of WTA responses are:

1. the participants care about the outcome;
2. the authority can enforce voters giving up the good;
3. the elicitation involves a yes or no vote on a single project; and
4. the probability that the proposed project is implemented is weakly monotonically increasing with the proportion of yes votes.

Suppose an individual in a group of M participants is presented with a single choice set s that asks for a vote between a project that is defined by attributes (a_s) and a cost (b_s) and the status quo. An individual's utility for the status quo is denoted by T_0 and the utility of project s is given by $U_s = u(Y + b_s; a_s)$ where Y is income. An individual m's vote can be represented by V_m. The votes potentially influence the probability that the policymaker will implement a certain policy. A "policy function" maps the collection of votes along with the policymaker's preferences G to a set of probabilities P that a certain policy indexed by n is implemented, $F(V_1, V_2, ...V_M; G) \to P = (P_1, P_2, ..P_N)$.[29]

Each individual is voting between a project s and the status quo. The individual's true preference ranking for the project is T_s such that $T_s =$ yes if $U_s \geq U_0$ and $T_s =$ no otherwise. Given T_s and all the votes of other participants, V_{-m}; the expected utility of the participant is

$$EU_m(T_s, V_{-m}; G) = P_s(T_s, V_{-m}; G)U_s + [1 - P_s(T_s, V_{-m}; G)]U_0$$

We can show that voting truthfully is a dominant strategy by considering what happens when individuals deviate from truthful voting. Let Vs be a non-truthfully vote (Vs Ts such that Vs = yes if Ts = no or Vs = no if Ts = yes). The change in expected utility from a non-truthful vote is equal to

[29]Note that $0 \leq P_n \leq 1\ \forall n$, $\sum_{n=1}^{N} P_n < 1$, and the probability that the status quo is maintained is $P_0 = 1 - \sum_{n=1}^{N} P_n$.

$$
\begin{aligned}
\Delta EU_m &= EU(V_s, V_{-m}; G) - EU(T_s, V_{-m}; G) \\
&= [P_s(V_s, V_{-m}; G) - P_s(T_s, V_{-m}; G)] (U_s - U_0)
\end{aligned}
$$

First, consider the case that $T_s =$ yes such that $U_s - U_0 \geq 0$. This implies that $V_s =$ no and $P_s(V_s, V_{-m}; G) - P_s(T_s, V_{-m}; G) \leq 0$ because of the monotonicity condition. Consequently, $\Delta EU \leq 0$. Conversely, if $T_s =$ no such that $U_s - U_0 \leq 0$, then $P_s(V_s, V_{-m}; G) - P_s(T_s, V_{-m}; G) \geq 0$ and once again $\Delta EU \leq 0$. Thus, deviating from one's true preferences cannot increase the participant's expected utility and this result holds regardless of the votes of others.

Appendix 3.B Robustness checks of public good WTA experiment results

Table 3.B.1: Public good experiment participant characteristics across treatments

	Consequential	Hypothetical	Real	Total
Age (years)	**26.4**	**26.7**	**25.9**	**26.4**
Hypothetical	1.00	-	-	
Real	1.00	1.00	-	
Male (%)	**39**	**44**	**40**	**42**
Hypothetical	0.68	-	-	
Real	1.00	0.76	-	
Volunteer (%)	**67**	**54**	**68**	**62**
Hypothetical	0.17	-	-	
Real	0.95	0.15	-	
Work part-time (%)	**43**	**32**	**44**	**39**
Hypothetical	0.24	-	-	
Real	1.00	0.2	-	
Work full-time (%)	**21**	**20**	**13**	**18**
Hypothetical	1.00	-	-	
Real	0.32	0.35	-	
Number of participants	**61**	**79**	**62**	**202**

Notes: Bolded numbers represent mean values for the continous variable age and the percentage of respondents for the other variables. The unbolded numbers report the p-value of a pair-wise t-test for the continuous variable age and of a pair-wise Chi-Squared test for the other variables between the different treatment samples.

Table 3.B.2: Probit regression with participant control characteristics

	Hypothesis 1	Hypothesis 2	Hypothesis 3
Hypothetical treatment	-0.472**		-0.562***
	(0.224)		(0.189)
Consequential treatment		0.105	
		(0.242)	
Age	-0.015	0.03	-0.003
	(0.016)	(0.024)	(0.014)
Male	-0.428*	-0.577**	-0.417**
	(0.224)	(0.254)	(0.191)
Volunteer	0.017	-0.11	-0.182
	(0.225)	(0.265)	(0.194)
Work part-time	-0.37	-0.44	-0.307
	(0.249)	(0.276)	(0.206)
Work full-time	-0.058	-0.868**	-0.322
	(0.298)	(0.386)	(0.26)
Constant	1.103**	0.258	0.993**
	(0.502)	(0.636)	(0.411)
Treatment included			
Hypothetical	Y	N	Y
Consequential	Y	Y	Y
Real	N	Y	Y
Observations	140	123	202
Log Likelihood	-90.99	-73.24	-128.7

Note: *p<0.1; **p<0.05; ***p<0.01

Assessing Alternative Probit Model Specifications

To assess whether there is heteroscedasticity with respect to the impacts of the treatment dummy variables in the probit models for hypothesis 1, 2 and 3, we conduct a series of likelihood ratio tests with different probit model specifications. Model 1 only includes the significant covariates from Table 3.B.2: male and working status (part-time or full-time work). Model 2 adds the treatment dummy variables (Hypothetical treatment for Hypothesis 1 and 3 and Consequential treatment for Hypothesis 2). Model 3 builds on Model 2 and is the standard heteroskedastic probit formulation where the treatment variable affects both the mean and variance of the underlying willingness to accept distribution. We conduct three likelihood ratio tests for each hypothesis.

For hypothesis 1, adding the hypothetical treatment dummy (i.e. comparing Model 1 and 2) results in a $\chi^2(1)$ test statistic equal to 4.56 (p-value = 0.033) which is consistent with Table 3.2. Adding the heteroskedastic variance parameter (i.e. comparing Model 2 and 3) results in a $\chi^2(1)$ test statistic equal to 0.666 (p-value = 0.417). Adding both the treatment dummy and variance parameter results in a $\chi^2(1)$ test statistic equal to 5.22 (p-value = 0.073).

For hypothesis 2, all model comparisons test statistics suggest that we cannot reject the hypothesis that responses in the consequential and real treatments are similar. For hypothesis 3, the first likelihood ratio test yields a $\chi^2(1)$ test statistic equal to 8.28 (p-value = 0.004) which is consistent with the results in Table 3.2. The second likelihood ratio test assesses whether adding the variance parameter improves model fit and here we cannot reject the hypothesis that there is no model fit improvement when the variance parameter is included. The third likelihood ratio test compares the simple model with only covariates to the heteroskedastic specification with the hypothetical treatment dummy variable affecting the mean and variance of the underlying WTA distribution. We find a $\chi^2(2)$ test statistic of 9.53 (p-value: 0.008) suggesting the $p = 0$ treatment results in a different voting pattern compared to the $p > 0$ treatments, which is consistent with Hypothesis 3.

Table 3.B.3: Log likelihoods and likelihood ratio tests for probit model specifications

	Hypothesis 1	Hypothesis 2	Hypothesis 3
Probit model specification		Log Likelihood	
Model 1: Covariates only (male, work part-time, and work-full time)	-93.72	-74.35	-133.29
Model 2: Covariates plus treatment dummy variable[a]	-91.44	-74.25	-129.15
Model 3: Covariates plus treatment dummy variable plus heteroskedastic error term[b]	-91.12	-74.24	-128.52
Likelihood ratio tests		Chi-squared (χ^2) test statistic (p-value)	
Model 1 vs Model 2: $\chi^2(1)$	4.56 (0.033)	0.19 (0.664)	8.28 (0.004)
Model 2 vs Model 3: $\chi^2(1)$	0.66 (0.417)	0.04 (0.845)	1.25 (0.264)
Model 1 vs Model 3: $\chi^2(2)$	5.22 (0.073)	0.23 (0.893)	9.53 (0.008)
Treatments included			
Hypothetical	Y	N	Y
Consequential	Y	Y	Y
Real	N	Y	Y
Observations	140	123	202

Notes: [a] As in Table 3.2, for Hypothesis 1 and 3, a dummy variable is added for the Hypothetical treatment (i.e. p = 0). For Hypothesis 2, a dummy variable is included for the Consequential treatment (i.e 0 <p <1).
[b] Model 3 is the standard heteroskedastic probit formulation where the treatment variable affects both the mean and variance of the underlying willingness to accept distribution.

Appendix 3.C Robustness checks of private good WTA experiment results

Table 3.C.1: Private good experiment participant characteristics across treatments

	Consequential	Hypothetical	Price	Provision	Total
Age (average years)	**26**	**28.5**	**26.4**	**24.9**	**26.4**
Hypothetical	0.23	-	-	-	
Price	0.93	0.93	-	-	
Provision	0.93	0.08	0.93	-	
Male (%)	**46**	**45**	**43**	**30**	**42**
Hypothetical	1	-	-	-	
Price	0.87	0.99	-	-	
Provision	0.14	0.23	0.3	-	
Volunteer (%)	**57**	**62**	**47**	**87**	**62**
Hypothetical	0.76	-	-	-	
Price	0.39	0.22	-	-	
Provision	<0.01	0.01	<0.01	-	
Work part-time (%)	**34**	**36**	**39**	**48**	**39**
Hypothetical	1	-	-	-	
Price	0.73	0.93	-	-	
Provision	0.2	0.35	0.49	-	
Work full-time (%)	**17**	**21**	**25**	**11**	**18**
Hypothetical	0.74	-	-	-	
Price	0.45	0.92	-	-	
Provision	0.54	0.29	0.14	-	
Number of participants	**65**	**42**	**49**	**46**	**202**

Notes: Bolded numbers represent mean values for the continous variable age and the percentage of respondents for the other variables. The unbolded numbers report the p-value of a pair-wise t-test for the continuous variable age and of a pair-wise Chi-Squared test for the other variables between the different treatment samples.

Table 3.C.2: Random effects logit models with various acceptance thresholds

	Lower-bound threshold level for a 'yes' response	
Parameter	Not sure (50% chance)	Definitely yes (100% chance)
Constant	-4.90***	-8.89***
	(0.671)	(1.012)
Amount	0.38***	0.39***
	(0.040)	(0.037)
Hypothetical treatment	0.66	0.83
	(0.548)	(0.756)
Provision framing treatment	0.96*	1.58*
	(0.539)	(0.742)
Price framing treatment	0.2	-0.23
	(0.542)	(0.743)
Provision motivation	3.58***	1.01
	(0.867)	(1.288)
Price motivation	-1.99**	-3.10*
	(0.833)	(1.230)
Amount*Provision motivation	-0.18***	-0.04
	(0.048)	(0.049)
Amount*Price motivation	0.18**	0.17***
	(0.057)	(0.051)
Observations	1,612	1,612
Participants	202	202
Log Likelihood	-398.53	-471.2
AIC	817.1	962.3

Note: *p<0.1; **p<0.05; ***p<0.01. AIC = Akaike Information Criteria. For all models, the Consequential treatment is the omitted reference category

Chapter 4

Incorporating Stated Consequentiality Questions in Stated Preference Research

Although the idea of consequentiality has played a transformative role in how stated preference (SP) practitioners think about the validity of survey responses, there remain lingering issues with implementing these concepts. One approach is to elicit stated consequentiality measures, but a concern arises regarding how to address the potential endogeneity concerns of these measures in econometric models of choice behavior. This study uses data from a drinking water reliability survey to shed light on this issue. We use a split sample approach that varies the order of the valuation and consequentiality questions and find that this treatment has a substantial and important impact on consequentiality perceptions. Furthermore, while consequentiality perceptions are important determinants in naïve voting models without endogeneity controls, using several methods to address endogeneity including the special regressor approach, I find that consequentiality does not have a significant impact on voting once the endogeneity has been addressed. These results suggest that the new trend of including consequentiality follow-up questions in surveys may not be the panacea to SP validity issues. The study also provides the first estimates of the public benefits of reducing boil water advisories (BWAs). Results suggest that people are willing to pay for improved reliability in small communities (under 500 people) but not larger sized communities.

4.1 Introduction

The stated preference (SP) literature has been transformed by the focus on ensuring respondents perceive their answers to be consequential to mitigate the hypothetical nature of SP surveys.[1] Carson and Groves (2007) outline the theoretical result that respondents' answers can be expected to be truthful if the survey is viewed by the respondent as potentially influencing the policy outcome and there exists some probability that they will have to pay. They suggest that consequentiality can be used as the incentive compatibility mechanism to ensure respondents tell the truth.

Herriges et al. (2010) provide an early empirical application of incorporating consequentiality perceptions, where they provide survey respondents with direct quotes from policy makers on the future use of the survey results to test this consequentiality hypothesis. Using willingness-to-pay (WTP) estimates, Herriges et al. (2010) find empirical evidence consistent with this prediction. Other empirical evidence on the importance of consequentiality in survey design has been provided by Landry et al. (2007), Vossler et al. (2012), and Interis and Petrolia (2014). Although the idea of consequentiality has played a transformative role in how SP practitioners think about the validity of survey responses, there are a number of lingering issues with implementing this concept. One concern is how to assess whether respondents think the survey is consequential. The most common practice is to ask respondents (using a Likert scale) their perceptions of the extent to which the survey results will be used by policy makers or affect decision-making. With these responses in hand, a second concern relates to incorporating these responses into the valuation estimation equation, potentially raising issues of endogeneity. Endogeneity can arise as responses to these consequentiality questions likely suffer from measurement error and are likely to be driven by factors unobserved to the analyst that are related to responses to the valuation questions (Herriges et al., 2010).

The aim of this paper is to explore the potential endogeneity of consequentiality questions in the econometric modeling of SP data. While this issue has been explored previously

[1] Understanding if and why people respond differently in hypothetical and real contexts has long been a focus of SP research. In general, this issue has been studied in experiments using both real and hypothetical payment treatments, with any resulting difference in behaviour being attributed to hypothetical bias. If preferences collected in SP surveys do not represent how individuals actually behave in the real world; the use of SP value estimates in economic analysis is questioned. Neoclassical economic explanations of why respondents misrepresent their preferences in SP surveys have largely focused on strategic behaviour. As a result, the focus of survey design has been on ensuring that respondents have the incentive to truthfully report their preferences (Carson and Groves, 2011). On the other hand, behavioural economics has offered a diverse set of explanations for hypothetical bias such as the social context, the level of scrutiny, and restrictions on the time-horizon and choice set (Levitt and List, 2007; Carlsson, 2010).

by Herriges et al. (2010) and Groothuis et al. (2017), I implement a new approach for addressing endogeneity: the special regressor approach.[2] I compare this approach to the common naïve method of incorporating consequentiality in voting behaviour models, as well as to a bivariate probit model and to a control function approach. A second novel aspect of the study is that a split sample approach is used in the survey that varies the order of the consequentiality and valuation questions. The study uses data from a SP survey implemented to value the public benefits of reducing boil water advisories (BWAs) in the province of Alberta, Canada.

This paper makes two contributions to the environmental valuation literature. The first contribution is to the growing body of theoretical and empirical evidence on the importance of consequentiality. The results of the split sample suggest that the order of the valuation and consequentiality questions matter for self-reported perceived levels of consequentiality. Specifically, in our application the percent of respondents who find the survey inconsequential increases by over 68 percent if the consequentiality question is asked after the valuation question compared to before (22.3 percent compared to 13.3 percent). This result has important implications, as attempts to exogenously manipulate the levels of perceived consequentiality by varying the information provided to respondents on use of surveys by policymakers have generally not been successful.[3] Furthermore, the results suggest that once the endogeneity of consequentiality is addressed, its importance in determining voting behaviour is less clear. Both of these results raise important issues for survey design.

A second contribution of this paper is to provide the first empirical evidence regarding the economic benefits of reducing BWAs. While there is a large literature evaluating the economic benefits of improving drinking water reliability, as far as I know, there is a lack of empirical evidence quantifying the economic benefits of reducing BWAs. The results suggest Albertans in small communities (less than 500 people) have a positive WTP to reduce BWA, but this does not hold for larger sized communities. Determining the value of reducing BWAs is important for public policy because the Canadian government has committed a substantial amount of money to reducing BWAs.

[2]The special regressor approach has previously been applied in the environmental valuation literature by Lewbel et al. (2011) to control for the endogeneity of double bounded choice questions and by Riddel (2011) and Kalisa et al. (2016) to address the endogeneity of risk perceptions.

[3]Herriges et al. (2010) was the first to successfully implement this strategy but subsequent attempts have not been successful (Oehlmann and Meyerhoff, 2017; Czajkowski et al., 2015).

4.2 Boil water advisories in Alberta

Access to reliable, clean drinking water continues to be an issue in many regions of Canada. One of the most common types of water reliability issues are BWAs. BWAs are mainly issued in response to high levels of turbidity in the water, presence of harmful microbes such as E. coli bacteria, or equipment and process failures or issues. In Alberta, BWAs are issued by Alberta Health Services as preventative measures to protect public health from waterborne infectious agents that may be present in drinking water. If water is consumed without boiling when there is an advisory, serious health problems can arise ranging from moderate illness to, in very rare circumstances, death.[4]

Communities have varying chances of being placed under a BWA depending on their source of drinking water and the condition of their water treatment system. The vast majority of BWAs issued in Alberta are for smaller towns and First Nations communities. Between 2008 and 2013, approximately 60% of the communities facing BWAs have been First Nations communities.[5] Table 4.1 presents the annual number of BWAs for three different community sizes in Alberta from 2008 to 2013. The typical BWA in Alberta last for 9 days.

Table 4.1: Number of boil water advisories (BWAs) in Alberta 2008-2013, by population size of community

	Community size			
Year	Small (<500 residents)	Medium (500 to 50,000 residents)	Large (>50,000 residents)	Total
2008	47	4	0	51
2009	40	3	1	44
2010	52	3	0	55
2011	49	4	0	53
2012	50	4	1	55
2013	59	5	1	65
Average	49.5	3.8	0.5	54

While the number of BWAs in Alberta can be reduced with new investments in water treatment facilities and watershed management, the costs are significant. The federal

[4] As the name suggests, when a BWA is issued, the public is advised to boil their tap water for drinking, preparing food, beverages, ice cubes, washing fruits and vegetables and brushing teeth. The water should be brought to a rolling boil for 1 minute to kill all disease-causing organisms.

[5] Author's calculations using publicly available BWA data collected from Alberta Health Service (2017) and Health Canada (2011).

government has committed in 2016 to spending \$1.8 billion over five years to end long-term drinking water advisories in First Nation communities (Indigenous and Northern Affairs Canada, 2017). While the investment needs are significant, even the most modern drinking water treatment facilities can fail due to significant adverse weather events such as floods and storms. Thus, it is likely prohibitively expensive to completely eliminate the possibility of BWAs. Understanding the benefits the public places on reducing BWAs is an important component of determining how much to invest in drinking water reliability.

4.3 Previous research on potential endogeneity of consequentiality questions

Several papers have explored the potential endogeneity of consequentiality responses in SP surveys. The first paper to tackle this problem head on was Herriges et al. (2010), exploring whether the consequentiality perceptions of respondents affect their WTP for improvements in lake water quality in Iowa. The source of endogeneity examined by Herriges et al. (2010) is the unobserved confounding problem - respondents who state a high degree of consequentiality may do so in part because they place a high value to the proposed environmental improvement programs. To assess and ultimately address these concerns, a split sample approach was implemented in which half of surveys included a letter from a government official stating that the information from these surveys will be used in the decision making process. The empirical results suggest that respondents' perceived degree of consequentiality is positively affected by the presentation of this letter. Using this exogenous information treatment, the causal impact of consequentiality perceptions on WTP is estimated using a Bayesian treatment effect model. The empirical findings are consistent with the "knife-edge" theoretical result of Carson and Groves (2007) which states that respondents who perceive their responses to have a positive probability of being taken into account by policy makers have similar WTP distributions while respondents who perceive the survey to be purely inconsequential have no incentive to respond truthfully and may have different WTP distributions.

While the findings of Herriges et al. (2010) suggest that the endogeneity of consequentiality questions is an important consideration in econometric modeling, three more recent papers find the opposite result. Vossler et al. (2012) test for the endogeneity of consequentiality questions using socio-demographic indicators such as occupation status, gender, and age as instruments. Using a generalized method of moments over-identification test, they fail to reject the hypothesis that the consequentiality interaction terms included in their model are jointly exogenous. Vossler and Watson (2013) use the same set of socio-demographic variables as instruments and tentatively conclude there is little empirical evidence for endogeneity of consequentiality responses, although the authors admit the instruments are weak. Finally, Interis and Petrolia (2014) use a similar set of socio-demographic instruments to test for endogeneity using a two-step instrumental variable (IV) probit model. The results suggest that the null hypothesis that consequentiality is exogenous cannot be rejected.[6]

Groothuis et al. (2017) use a bivariate probit approach to address endogeneity of perceived consequentiality responses. The correlation coefficient between the two error terms is estimated to be negative and significant, suggesting that perceived consequentiality is endogenous and that the unobserved characteristics increase perceived levels of consequentiality and decrease the likelihood of voting for the program.

Another proposed approach to dealing with endogeneity of consequentiality is the Hybrid Choice Model (Czajkowski et al., 2015). However, Budziski and Czajkowski (2017) demonstrate through simulations that Hybrid Choice Models do not eliminate the bias in estimated coefficients of endogenous variables.

Table 4.2 summarises the previous research investigating the potential endogeneity of consequentiality questions in SP research. The mixed evidence found in these studies regarding the endogeneity of consequentiality questions can likely be explained by the different empirical applications, the use of different instruments and modeling approaches. The information treatment used by Herriges et al. (2010) is perhaps the most convincing instrument for addressing the endogeneity of consequentiality questions. Subsequent attempts to replicate these effects using different forms of consequentiality scripts, however, have been largely unsuccessful (Oehlmann and Meyerhoff, 2017; Czajkowski et al., 2015). Socio-demographic characteristics of respondents are generally not strongly correlated with perceived consequentiality and therefore their usefulness as instruments is limited. Thus the elusive search for a robust instrument for perceived consequentiality responses in voting behaviour models continues.

[6]However, in addition to weak instruments, the IV probit model produces inconsistent estimates if the endogenous variable is not continuous (Lewbel, 2014), such as the case with consequentiality perceptions reported on the Likert scale .

Table 4.2: Summary of literature that test for endogeneity of consequentiality responses

Study	Empirical Application	Approach	Instrumental Variable	Endogeneity Result
Herriges et al. (2010)	Lake water quality	Bayesian treatment effect model	Split sample treatment with letter from policy maker	Endogenous
Vossler et al. (2012)	Tree planting	GMM over-identification test	Socio-demographic indicators (gender, age, and income, occupation)	No evidence for endogeneity
Vossler and Watson (2013)	Preserve and restore natural areas	Not reported	Socio-demographic variables	No evidence for endogeneity
Interis and Petrolia (2014)	Coastal wetland restoration	Instrumental variable probit	Socio-demographic indicators (gender, age, minority, and income)	No evidence for endogeneity
Czajkowski et al. (2015)	Discounted theatre tickets	Control function	Not reported but noted it is not strong	No evidence for endogeneity
Groothuis et al. (2017)	Water conservation	Bivariate probit	Local control*	Endogenous

* The specific variable is a dummy variable for whether the respondent agreed or strongly agreed with the statement: "Local public officials (city/county) should have the final authority to make decisions about how our water supply is managed".

4.4 Survey data

4.4.1 Survey development and design

I use data from a SP survey on drinking water reliability in Alberta. The survey was designed using three focus groups conducted in the spring of 2014 in Edmonton (n=10), Calgary (n=8), and Okotoks (n=11), as well as a pilot test involving 155 respondents. The survey was divided into three sections and included 52 questions.[7] The first part of the survey introduced the respondents to drinking water reliability issues in Alberta and collected information on past experiences and future risk perceptions of short and long term water outages and BWAs. The second part of the survey included the BWA valuation question that is the focus of the current paper. Before the valuation question, respondents were provided with additional information on the frequencies of BWAs by community size over the past 5 years to inform the current situation.

[7]The full survey instrument is included in Supplementary Material E.

Respondents were then told about a proposed program that would reduce the annual number of BWAs in Alberta. The scope of the BWA reduction varied by small (under 500 residents), medium (between 500 and 50,000 residents) and large (over 50,000 residents) community sizes. The payment vehicle was described as additional income taxes collected over the next 10 years.

The valuation task used a single binary choice (SBC) question. While collecting less information per respondent compared to repeated multinomial-choice question formats, the SBC has several key advantages. First, it closely mimics a real referendum vote which, combined with the clear public goods framing of the program, enhances incentive compatibility (Carson and Groves, 2007). Second, it avoids the issues of ordering effects, learning, strategic behaviour, and other potential biases associated with repeated choice questions (Johnston et al., 2017).

The choice task was described as a referendum vote between the current situation and the proposed program. Table 4.3 provides a description of the attributes and their respective levels that are used in the SBC question. The SBC employed a D-efficient design consisting of 31 final choice tasks.[8] The number of BWAs for the current situation is informed by historical data presented in Table 4.1 and includes 50 BWAs in small communities, 4 BWAs in medium communities, and 1 BWA in large communities. The attribute levels in the proposed program range from 5-50 for BWAs in small communities, 1-4 BWAs in medium communities, and 0-1 BWAs in large communities. In addition to reducing BWAs, the proposed program also included an attribute that specified the specific method used to improve reliability. The first reliability improvement method was investments in traditional drinking water treatment systems (i.e. grey infrastructure). The second improvement method was investments in watershed and forest management to reduce the potential for events such as forest fires to cause water reliability problems downstream (i.e. green infrastructure). The final section of the survey includes debriefing questions that collected socio-demographic information.

One other notable aspect of the survey is that a 2x2 design was used on two different treatments. First, a split sample treatment was implemented with half the surveys including the consequentiality question before the valuation question and half the surveys including the consequentiality question after.[9]

[8]The choice tasks were designed using NGENE. I removed one of the initial 32 choice tasks created by NGENE that only differed by the treatment method and cost between the status quo and program.

[9]The specific question is *To what extent do you believe that the voting results collected from you and other survey respondents will be taken into consideration by policy makers?* Respondents answered on a five-point likert scale from ‘Not taken into account’ to ‘Definitely taken into account’.

Table 4.3: Single binary choice question attributes and levels

Attribute	Definition	Levels[a]
BWA small	Annual number of BWAs in small communities with less than 500 residents	5,15,25,**50**
BWA medium	Annual number of BWAs in medium-sized communities with between 500 and 50,000 residents	1,2,3,**4**
BWA large	Annual number of BWAs in Large communities with more than 50,000 residents	0,**1**
Treatment method	Method used to improve reliability	Drinking water treatment system, Watershed and forest management
Tax Amount	Annual Increase in provincial income tax for 10 years	**$0**, $10, $50, $125, $250

Notes: BWA = boil water advisory. [a] Bolded text denotes the current situation level. The method to improve reliability for the current situation was described as the current system.

Second, half the surveys included a statement on the government agencies involved in the project. Specifically, at the beginning of the survey the following statement was included: "Partners in this project include Alberta Environment and Sustainable Resource Development and the Canadian Forest Service." The purpose of this, admittedly quite vague, statement was to act as an information treatment to enhance perceived consequentiality similar to Herriges et al. (2010).

4.4.2 Survey administration

An online pilot survey was administered to 155 respondents between January and February, 2015. The experimental design was updated based on the responses to the pilot survey using a D-efficient design. The final survey was administered online by an Edmonton based market research firm in March 2015. In addition to a representative sample of 1,000 Alberta residents, 250 additional responses from residents of rural communities were included because water reliability challenges are more common in rural communities.[10]

The survey also included a separate set of valuation questions focused on private home water outages and the order of these valuation questions was randomized. For the empirical analysis, I only include the 757 respondents who received the BWA valuation question

[10] The survey research firm does not disclose response rates for the survey. The completion rate was 60% for the 2,105 people who started the survey.

before the other valuation questions to avoid ordering issues.[11]

Descriptive statistics for these respondents are presented in Table 4.4. The average age of respondents was 49 years and 31% of respondents had household incomes over $150,000 per year. A further 12% of respondents did not answer the income question. In the sample, 37% of respondents had attended some or completed college and 45% of respondents had completed at least some university. Comparing our sample with the 2011 Canadian census results for Alberta, the sample contains a similar gender mix compared to the general population (51% versus 50% female) but a higher proportion of people over the age of 50 (56% compared to 44% in the general population).[12] The sample also included 31% of respondents with household incomes greater than $150,000 per year compared to only 18% in the broader Alberta population. Thus our sample is older and has higher incomes compared to the general population.

Table 4.4: Socio-demographic summary statistics

Variable	Description	Mean	Std. Dev.	Min	Max
Female	Respondent is a female	0.51	0.50	0	1
College	Completed at least some college or technical school	0.37	0.48	0	1
University	Completed at least some university	0.45	0.50	0	1
High income	Household income greater than $150,000	0.31	0.46	0	1
Income missing	Did not answer income question	0.12	0.32	0	1
Age index	Age index is calculated by dividing the age of each respondent by the mean age of the sample.	1.01	0.30	0.54	2.31
Live in large community	Lives in community with more than 50,000 people	0.48	0.50	0	1
Live in medium community	Lives in community with between 500 and 50,000 people	0.18	0.38	0	1
Live in small community	Lives in community with less than 500 people	0.34	0.47	0	1
Vote for program	Voted for the proposed water reliability program	0.49	0.50	0	1

Notes: Sample size: n = 757

[11] A total of 769 respondents fit this criterion, but 5 respondents did not answer the voting question, 2 respondents did not answer the consequentiality question, and 5 respondents did not answer the age question.

[12] These percentages are computed excluding census data for people under the age of 18 because they were not eligible for the survey. Census data collected from (Statistics Canada, 2011).

4.5 Econometric model

The econometric analysis of stated choices is grounded in McFadden's (1973) random utility model (RUM). In a single binary choice setting, respondents are asked to choose between two alternatives: a status quo alternative and a program alternative. Both alternatives are characterized by various attributes. The RUM is based on the idea that an individual chooses the alternative that yields the highest expected utility between the two choices.

I start with the following binary decision model:

$$D_i = I(X_i'\beta + C_i'\delta + B_i\alpha + \varepsilon_i > 0) \tag{4.1}$$

where $I(.)$ represents an indicator function equal to one when the argument is true and zero otherwise, D_i is the yes/no answer to the valuation question for respondent i, C_i is the response to the perceived consequentiality question, B_i is the tax amount presented to the respondent, X_i is a vector of exogenous regressors, and ε_i is the error term. If C_i is assumed to be exogenous, then the parameters of the model can be estimated using a probit model and WTP can be derived in the conventional fashion.

4.5.1 Bivariate probit model

The traditional approaches to handling potentially endogenous binary variables in non-linear models are based on maximum likelihood methods (Wooldridge, 2015). In the case of a binary main equation such as a binary choice model, a common specification is the bivariate probit model. For the bivariate probit model, a second equation is introduced that describes the perceived consequentiality response. This equation can be written as

$$C_i = I(X_i'\gamma + Z_i'\theta + \omega_i > 0) \tag{4.2}$$

where Z_i is a vector of instruments, ω_i is an error term, and γ and θ are parameters to be estimated. The robustness of the bivariate probit model relies on two key assumptions. First, Equation (4.2) is correctly specified. Second, the two error terms $(\varepsilon_i, \omega_i)$ of the Equation (4.1) and Equation (4.2) follow a joint normal distribution. If these two conditions hold, than maximum likelihood procedures can consistently estimate the parameters β, δ, and α in Equation (4.1).

4.5.2 Control function approach

Control function approaches use a similar set-up to the bivariate probit model, but estimation is done through two steps (Wooldridge, 2015; Lloyd-Smith et al., 2016). In our context, the estimated residuals from Equation (4.2) are included in the probit model to act as controls for the endogeneity of the consequentiality variables. There is some confusion in the literature on whether control function approaches can be used in nonlinear models with discrete endogenous variables. Bontemps and Nauges (2016) and Kalisa et al. (2016) both state that control function approaches are inconsistent. This is generally true for the commonly used IV probit approach which uses a linear regression in the first stage (i.e. Equation (4.2)). However, control function approaches need not use a linear functional form for the first stage. As shown by Wooldridge (2015), if a probit model is used in the first stage, combined with including the generalized residuals in the second stage, the control function approach is no more or less robust than the bivariate probit model.

The generalized residual can be computed from the first stage probit model as

$$r_i = C_i \lambda(X_i'\hat{\gamma} + Z_i'\hat{\theta}) - (1 - C_i)\lambda(-(X_i'\hat{\gamma} + Z_i'\hat{\theta}))$$

where λ is the inverse Mills ratio. As noted by Wooldridge (2015), the key distributional assumption needed is to assume that ε_i depends on (Z_i, D_i) only through r_i in the conditional distribution sense. The control function approach shares the same probit reduced form for C_i but uses different assumptions about the conditional distribution.[13] Thus, similar to the bivariate probit model, it relies on the first stage equation being correctly specified, but not the jointly normal assumption regarding the error terms.

4.5.3 Special regressor approach

A different approach to controlling for endogenous regressors is provided by Lewbel (2000). He introduced a simple multi-step estimator for the scaled probit model with a

[13] The term generalized residual comes from the fact it has a mean of zero conditional on $(X_i'\hat{\gamma} + Z_i'\hat{\theta})$ (Wooldridge, 2015).

very exogenous regressor that is useful given our discrete choice setting. To operationalize the approach, I can re-write Equation (4.1) so that the special regressor's coefficient is normalized to one:

$$\widetilde{D_i} = I(X_i'\widetilde{\beta} + C_i'\widetilde{\delta} + B + \varepsilon_i > 0) \tag{4.3}$$

The special regressor, B, must have the following properties:

1. B is additively separable with respect to the model error ε_i,
2. B is independent of the model error, ε_i, conditional on the set of regressors (i.e. B is exogenous),
3. $E(D|X, C, B)$ increases with B, and
4. The conditional distribution of B given X and C is continuous and has a large support.

In our application, I use the randomly assigned tax amount presented to respondents as the special regressor which has been used in previous research (Lewbel et al., 2011; Riddel, 2011; Kalisa et al., 2016). Property 1 is satisfied if I use a linear functional form of the indirect utility function as is common in the discrete choice literature. Property 2 is satisfied because the specific tax amounts are selected by the researcher and randomly presented to respondents as part of the experimental design.[14] We would expect the probability of voting yes to the proposed program decreases with the tax amount and therefore to ensure Property 3 holds, I use the negative of the tax amount as the special regressor. Property 4 is a common assumption in semi-parametric binary choice models (Horowitz, 1992) and in our application implies that the support of the distribution of WTP is large relative to the model error, ε (Kalisa et al., 2016).

Estimation in the special regressor approach proceeds in four steps (Riddel, 2011):

[14] One potential limitation of using the tax amount as a special regressor to control for the endogeneity of perceived consequentiality is if the level of consequentiality is affected by the tax amount presented to respondents. Groothuis et al. (2017) provide some empirical evidence that as the tax amount given to respondents increase, the levels of perceived consequentiality decrease. In our application, I test for this effect by including the tax amount in the perceived consequentiality equation and do not find a significant effect. Furthermore, because of the split-sample decision that switched the order of the consequentiality and valuation questions, I can also use the subsample that stated their perceived consequentiality levels before being presented with the choice task to ensure the exogeneity of the tax amount.

Step 0: Create $\widetilde{B} = B - E[B]$ to allow $\widetilde{B}$ to take on a range of positive and negative values.

Step 1: Estimate the equation $\widetilde{B} = \vartheta X_i + \phi C_i + \mu_i$ using a linear regression and save the residuals $\hat{\mu_i}$.

Step 2: Compute the non-parametric kernel estimator of the density f of $\hat{\mu_i}$, $f(\hat{\mu_i})$, for each $\hat{\mu_i}$ using the following equation:

$$f(\hat{\mu_i}) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{\hat{\mu_j} - \mu}{h}\right),$$

and compute the estimates $\hat{f_i} = \hat{f}(\hat{\mu_i})$. For this step, I require a choice of kernel $K(.)$ and a bandwidth h.[15]

Step 3: Construct $\hat{T_i}$ for each observation i as:

$$\hat{T_i} = \frac{D_i - I(\widetilde{B_i} \geq 0)}{\hat{f_i}},$$

where $I(\widetilde{B_i} \geq 0)$ is an indicator function equal to 1 if $\widetilde{B_i} \geq 0$ and 0 otherwise.[16]

Step 4: Estimate the choice-model parameters of the scaled probit $(\widetilde{\beta}, \widetilde{\delta})$ using a two-stage least squares regression of $\hat{T_i}$ on X_i and C_i using instruments Z_i.

These four steps comprise the special regressor approach to recovering the parameters of the scaled probit model in Equation (4.3) that are used in this paper.[17] The coefficient on the tax amount is interpreted as the marginal utility of income and in the current application is normalized to one. Therefore, the coefficients on the other model variables $(\widetilde{\beta}, \widetilde{\delta})$ can be interpreted as WTP.[18]

[15] An alternative approach for this step is to use the ordered data estimator proposed by Lewbel and Schennach (2007).

[16] As noted by Bontemps and Nauges (2016), the denominator, $\hat{f_i}$, can take on very small values which implied extremely large $\hat{T_i}$ absolute values. This can induce large standard errors in the two-stage least squares regression in Step 4. Consequently, Lewbel (2014) recommends some form of trimming or winsorising to remove these extreme values. For our main specification, I use a 5% winsorisation specification.

[17] The special regressor approach is implemented in Stata using Baum's 2012 *sspecialreg* routine.

[18] This normalization has no impact on marginal effects and estimated WTP because the parameters of probit models are only identified up to the location and scale. Specifically, the parameters are related by $(\beta, \delta, \alpha) = (\widetilde{\beta}, \widetilde{\delta}, 1)\frac{1}{\sigma^{\varepsilon}}$ where σ^{ε} is the root mean-square error of the regression (Riddel, 2011).

4.5.4 Instrumental variables

The bivariate probit and control function approaches require at least one instrumental variable, while the special regressor approach can use an instrumental variable in Step 4. The main instrument I consider is the voting and consequentiality question order.[19] As illustrated in the next section, this variable has a significant effect on levels of perceived consequentiality. Because the survey versions were randomized, I do not expect this variable to affect voting behaviour except through its impact on consequentiality.

4.5.5 Weighting

Because respondents living in small communities were oversampled, I need account for this oversampling in the estimation. In the representative sample, 62.5% of respondents live in large communities, 23% live in medium communities, and 14.5% live in small communities. As shown in Table 4.4, respondents living in small communities constitute 34% of the sample used in the analysis, or 256 respondents out of 757. I use bootstrap sampling to repeatedly sample 85 respondents from these 256 people living in small communities to correct for the oversampling. This sampling is done 500 times for each of the models presented in the paper. Appendix 4.B provides the full set of empirical results without weighting.

4.6 Results

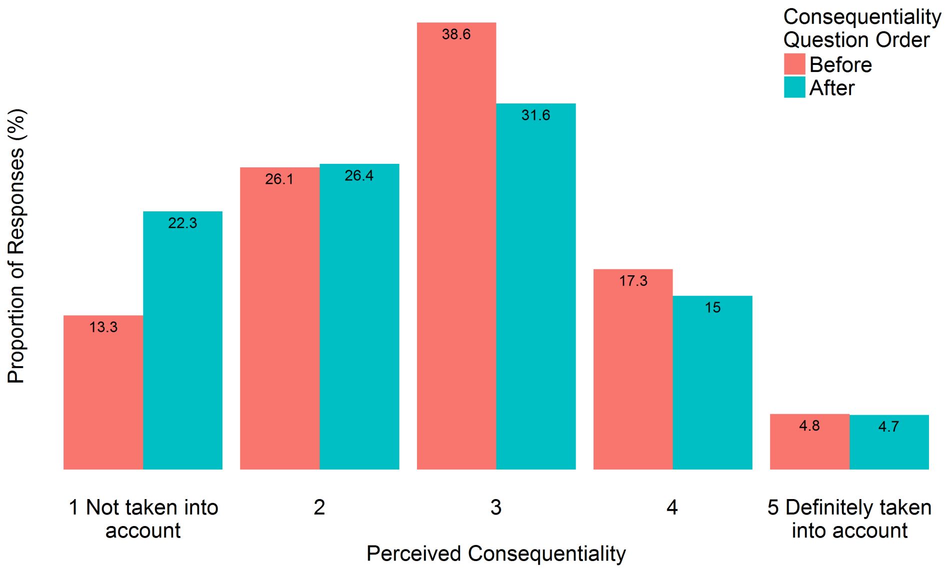
4.6.1 Effect of question order on perceived consequentiality

Figure 4.1 provides a comparison of the distributions of perceived consequentiality responses for the two survey versions that varied the question orders. This figure provides

[19] As a robustness check, I also consider an alternative instrumental variable - whether the respondent would vote for the incumbent government's political party. This political support dummy variable equals 1 if the respondent indicated they would vote for the incumbent political party in the upcoming provincial election and 0 otherwise. At the time of the survey, the political party holding power in the provincial government had been there for 43 years and I expect party supporters to be more inclined to perceive that their responses would be taken into account by decision makers. However, there are reasonable doubts about the strict exogeneity of this instrument as these party supporters may be more or less inclined to favour the proposed program for reasons besides perceived consequentiality. At the very least, this dummy variable is insignificant in the voting probit model. The results obtained using this instrument are similar to the results presented in the paper and the instruments pass the Sargan-Hansen test of over-identification.

preliminary evidence regarding the importance of the question ordering. A more formal test of distribution differences can be conducted using Pearson's Chi-squared test. The returned test statistic is 11.9 (p-value = 0.018) and thus I can reject the null hypothesis at the 5% significance level that the consequentiality responses are independent of the ordering treatment.[20] The most notable result in Figure 4.1 is that the proportion of respondents indicating that their voting results will 'not be taken into account' by policy makers jumps from 13.3% to 22.3% if the consequentiality question is after, rather than before, the valuation question. The results suggest that the ordering of the consequentiality question has a marked impact on the least consequential responses.

Figure 4.1: Ordering Effect on Perceived Consequentiality



Notes: To isolate the effect of ordering, the figure only includes the first consequentiality response for each respondent. The 'before' treatment includes 375 respondents and the 'after' treatment includes 382 respondents

4.6.2 Determinants of perceived policy consequentiality

This section presents a more formal analysis of factors that impact perceived policy consequentiality. We convert the consequentiality perceptions variable into a 'knife-edge' dummy variable that takes a value of 0 if the respondent stated that his/her responses would "not be taken into account" and the value of 1 otherwise (Oehlmann and Meyerhoff,

[20]While randomization should control for the influence of socio-demographic differences, a formal test of the two split samples provided in Table 4.A.1 suggest some statistically significant differences regarding the size of communities where the respondent lives. The next section uses regression models to control for any systematic differences between the two samples.

2017).[21] Table 5 reports the results of various probit model specifications with the the binary consequentiality variable as the dependent variable. The first column includes the "consequentiality question after" treatment dummy variable which has a negative and significant effect on the perceived consequentiality which is consistent with Figure 4.1. In contrast to Groothuis et al. (2017), I do not find a significant effect of the tax amount presented to respondents on perceived consequentiality. Similar to Vossler and Watson (2013) and Oehlmann and Meyerhoff (2017), there is a rather weak relationship with perceived consequentiality and socio-demographic variables. The partner information variable represents whether the survey version included the statement on government partner agencies involved in the research. The coefficient is not statistically insignificant which corroborates the findings of Oehlmann and Meyerhoff (2017) and Czajkowski et al. (2015) of the difficulty in inducing consequentiality perceptions.

To investigate the impact of the valuation question attributes on perceived consequentiality, the second column of Table 4.5 shows the results using the subsample that received the consequentiality question after the valuation question. The BWA reduction for medium communities has a negative and significant impact on perceived consequentiality while the other program attribute coefficients are insignificant. These results suggest that the reason for the difference shown in Figure 4.1 can be partly explained by the valuation question attributes. Column three of Table 4.5 uses the sample of respondents that received the consequentiality question before the valuation question. As expected, none of the program attributes have a significant effect on perceived levels of consequentiality.

4.6.3 Effects of perceived consequentiality on voting behaviour

Table 4.6 shows the results of the probit model of water management referendum voting without addressing the potential endogeneity of consequentiality responses. For all these models, I convert the BWA attribute variables to represent reductions from the status quo level. The first column is the base model without any controls for consequentiality. In this specification, I find that respondents prefer programs with more reductions in the number of BWAs in small communities, but no clear preferences for reducing BWAs in medium or large communities. Using a conventional drinking water treatment system or watershed and forest management has no impact on voting behaviour. The coefficient on the tax amount is negative and significant. In terms of socio-demographic variables, gender has

[21] I also estimated an ordered probit model using the full five point scale and find that the ordering treatment has a significant effect on consequentiality perceptions. These results are available upon request.

Table 4.5: Probit model of perceived consequentiality

	All Respondents	Consequentiality After Sample	Consequentiality Before Sample
BWA small	0.0015	-0.0006	-0.0011
	(0.00431)	(0.00583)	(0.00677)
BWA medium	-0.0791	-0.2040**	0.0412
	(0.0637)	(0.0850)	(0.107)
BWA large	-0.1030	-0.1980	0.1523
	(0.134)	(0.178)	(0.238)
Forest management treatment	0.0590	-0.1996	0.2821
	(0.132)	(0.196)	(0.229)
Tax amount	0.000186	0.000990	-0.00138
	(0.000753)	(0.00104)	(0.00123)
Female	0.120	0.134	-0.0448
	(0.123)	(0.191)	(0.231)
College	-0.144	-0.311	-0.0591
	(0.203)	(0.659)	(0.584)
University	-0.0433	-0.105	-0.221
	(0.515)	(1.128)	(0.753)
High income	-0.218	-0.285	0.0738
	(0.498)	(1.086)	(0.743)
Income missing	-0.293	-0.471	-0.116
	(0.500)	(1.075)	(0.718)
Age index	-0.307	0.0115	-0.864**
	(0.194)	(0.328)	(0.336)
Live in large community	0.195	0.178	0.135
	(0.138)	(0.187)	(0.244)
Consequentiality question after	-0.340***		
	(0.132)		
Partner information[a]	0.0416	0.329*	-0.203
	(0.130)	(0.187)	(0.229)
Constant	1.510***	1.261*	2.109***
	(0.367)	(0.751)	(0.793)
Observations	586	297	288
Log likelihood	-261.5	-145.3	-102.6
Chi-squared	20.77	14.50	13.74

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. [a] A dummy variable equal to 1 if the information treatment for the government agency partner information is presented at beginning of survey.

a significant impact on voting behaviour with females having greater preferences for the program compared to males.

The second column includes a perceived consequentiality variable using the same dummy variable as the previous section (Vossler and Watson, 2013; Groothuis et al., 2017). The coefficient on the consequentiality variable is positive and significant suggesting that respondents that perceive the survey to be at least somewhat consequential are more likely to vote for the program. The other explanatory variables remain relatively stable. The last two columns report the results using the consequentiality dummy variable to split the sample. The third column includes the consequential respondents (C>1) and the program attribute results are similar to the two previous columns. Using the consequential subsample, high-income respondents are more likely to accept the program compared to low-income respondents. The last column includes the inconsequential sample (C=0). While acknowledging the small sample size (n=107). I find that the coefficients on all the BWA reduction attributes are not statistically significant at the 10% level. While these models highlight the importance of incorporating consequentiality, they assume that perceived consequentiality is exogenous in these voting models.

4.6.4 Controlling for endogeneity of consequentiality perceptions

Table 4.7 presents the results of the three models that address the potential endogeneity of perceived consequentiality, along with the naïve consequentiality dummy probit model from Table 4.6. For the bivariate probit specification, the results are very similar to the probit model. This finding can be explained by the fact that the likelihood-ratio test that the two probit models are unrelated is not rejected at the 10% significance level. This result suggests that consequentiality is not in fact endogenous. The main difference between the probit and bivariate probit model is that while the coefficient on consequentiality increases in magnitude in the bivariate probit model it is not statistically significant due to the larger standard errors. A similar result is found using the control function approach. The estimated coefficient on the included residual is not statistically significantly different from zero which suggests that consequentiality is not endogenous. Nonetheless, the coefficient on the consequentiality dummy variable is not statistically significant different from zero due to the large standard error using the control function approach.

The special regressor approach results are presented in the fourth column of Table 4.7. These estimates are given by a two stage least squares regression linked to Step 4 of the

Table 4.6: Naive probit model of water management referendum voting

	Base Model	Consequentiality Dummy Model	Consequential Model (C>1=1)	Inconsequential Model (C=0)
BWA small	0.0112***	0.00834**	0.00620*	0.0191
	(0.00342)	(0.00334)	(0.00370)	(0.0117)
BWA medium	0.0152	0.0350	0.0402	-0.0528
	(0.0454)	(0.0518)	(0.0537)	(0.150)
BWA large	0.152	0.115	0.162	-0.178
	(0.116)	(0.111)	(0.120)	(0.339)
Forest management treatment	-0.107	-0.117	-0.0958	-0.295
	(0.116)	(0.110)	(0.122)	(0.362)
Tax amount	-0.00287***	-0.00253***	-0.00305***	-0.00425*
	(0.000630)	(0.000624)	(0.000692)	(0.00233)
Female	0.363***	0.326***	0.376***	0.126
	(0.110)	(0.117)	(0.117)	(0.377)
College	0.246	0.151	0.244	0.370
	(0.162)	(0.161)	(0.181)	(0.615)
University	-0.0424	0.0817	-0.293	1.117
	(0.309)	(0.312)	(0.363)	(3.301)
High income	0.383	0.155	0.686**	-0.901
	(0.304)	(0.310)	(0.343)	(3.238)
Income missing	0.576*	0.411	0.764**	-0.821
	(0.303)	(0.320)	(0.343)	(3.224)
Age index	-0.0929	-0.0401	0.0779	-1.260*
	(0.197)	(0.197)	(0.213)	(0.705)
Live in large community	0.0678	0.0431	0.0553	0.146
	(0.111)	(0.118)	(0.131)	(0.362)
Consequentiality Dummy (C>1=1)		0.421***		
		(0.145)		
Constant	-0.445	-0.706**	-0.474	0.847
	(0.302)	(0.331)	(0.319)	(1.236)
Observations	586	586	484	107
Log likelihood	-381.4	-383.5	-314.7	-59.38
Chi-squared	43.82	37.45	38.17	11.08

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

special repressor approach outlined in the previous section. The coefficient of the special regressor B is not shown here because it is normalized to one. Therefore, the other coefficients can be interpreted as WTP. I withhold discussion of the WTP coefficients until the next section, but we can note that the coefficient for the consequentiality variable is not statistically significantly different from zero.

In sum, the results of the models that address potential endogeneity of consequentiality responses consistently suggest that perceived consequentiality is not a statistically significant determinant of voting behaviour. This result is in contrast to the naïve probit model where perceived consequentiality is an important determinant of voting behaviour. However, I do not find evidence to suggest that consequentiality is in fact endogenous.

4.6.5 Willingness-to-pay to reduce boil water advisories

Table 4.8 summarises the WTP estimates for a single BWA reduction in the three different community sizes. The WTP results for the probit model with the whole sample, probit model with consequential respondents, bivariate probit model, and control function models are quite similar. The general public has a positive WTP of \$2.73 to \$3.47 per household per year for 10 years to reduce a single BWA in a small community. While the estimated WTP generally increase the reductions in BWAs in medium and large communities, these estimates are not statistically significantly different from zero. Using the special regressor approach, the WTP for a BWA reduction in a small community is estimated to be \$1.56, substantially less than the other models.

4.7 Conclusion

The perceived importance of consequentiality perceptions of respondents in SP surveys has seen a rapid rise in recent years. Ensuring survey responses are consequential is now a 'best practice' in SP work (Johnston et al., 2017). And yet questions remain on the best way to elicit consequentiality perceptions of respondents and, more importantly, on how to incorporate this information in the econometric analyses of SP data. In all survey applications eliciting consequentiality perceptions to date, the consequentiality question is posed after the valuation question. This study provides the first empirical evidence that the order of the consequentiality and valuation question matters for consequentiality perceptions. Specifically, the number of respondents who view the survey as inconsequential increases 68 percent if the valuation question is posed first (13.3 percent

Table 4.7: Alternative methods to address endogeneity

	Probit	Bivariate Probit	Control Function	Special Regressor[a]
BWA small	0.00834**	0.0107***	0.00773**	1.562**
	(0.00334)	(0.00329)	(0.00387)	(0.608)
BWA medium	0.0350	0.0521	0.0174	-1.193
	(0.0518)	(0.0629)	(0.0698)	(15.63)
BWA large	0.115	0.104	0.0815	-5.396
	(0.111)	(0.112)	(0.125)	(21.65)
Forest management treatment	-0.117	-0.182*	-0.103	-18.27
	(0.110)	(0.107)	(0.120)	(20.90)
Tax amount	-0.00253***	-0.00308***	-0.00256***	-
	(0.000624)	(0.000718)	(0.000692)	-
Female	0.326***	0.337***	0.336***	43.71**
	(0.117)	(0.109)	(0.128)	(20.31)
College	0.151	0.164	0.280	21.18
	(0.161)	(0.152)	(0.174)	(27.85)
University	0.0817	-0.0818	-0.125	-39.53
	(0.312)	(0.472)	(0.368)	(69.76)
High income	0.155	0.332	0.396	66.35
	(0.310)	(0.450)	(0.341)	(63.76)
Income missing	0.411	0.617	0.641*	77.02
	(0.320)	(0.449)	(0.356)	(62.54)
Age index	-0.0401	0.0102	-0.129	-1.951
	(0.197)	(0.220)	(0.287)	(45.59)
Live in large community	0.0431	0.0381	0.0383	-2.132
	(0.118)	(0.134)	(0.163)	(39.19)
Consequentiality Dummy (C>1=1)	0.421***	0.533	0.840	9.792
	(0.145)	(1.080)	(1.016)	(445.4)
Residual			0.283	
			(0.550)	
Constant	-0.706**	1.835***	-0.603	-51.18
	(0.331)	(0.387)	(0.379)	(403.8)
ρ		-0.0715		
		(9.408)		
N	586	586	586	586

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. [a] The coefficient of the special regressor B is not shown here because it is normalized to one. Therefore, the other coefficients can be interpreted as WTP. The estimates are given by a two stage least squares regression linked to Step 4 of the special repressor approach outlined in the previous section.

Table 4.8: Mean annual willingness-to-pay for 10 years to reduce boil water advisories

	Probit Model (Full sample)	Consequential Probit Model (C>1=1)	Bivariate Probit Model	Control Function	Special Regressor
BWA small	\$3.30**	\$2.73*	\$3.47***	\$3.02*	1.56**
	(1.502)	(1.515)	(1.210)	(1.615)	(0.608)
BWA medium	\$13.84	32.87	16.95	6.82	-1.19
	(20.78)	(21.37)	(20.79)	(27.61)	(15.63)
BWA large	\$45.27	\$49.95	\$33.96	\$31.86	-5.40
	(43.93)	(46.29)	(37.18)	(49.51)	(21.65)
hline N	586	484	586	586	586

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

to 22.3 percent). This finding is important in light of the theoretical and empirical knife-edge results identified by Carson and Groves (2007) and documented by Herriges et al. (2010).

The question of why we might expect such a large difference in perceived levels of consequentiality remains an open question. The results of this paper suggest that a respondent's consequentiality beliefs are influenced by at least one of the program attributes and thus the specific program offered to respondents may matter. A somewhat speculative explanation for why the proportion of respondents who respond to the consequentiality question second find the survey to be more inconsequential is that going through the voting process may cause respondents to doubt how applicable the results will be for decision-makers. The online voting mode administered to respondents is quite different from the traditional way policy-makers have administered referendum surveys via the mail. An alternative explanation is that once the respondents have seen the program and voted, they have a greater probability of engaging in K-level thinking and acting strategically in responding to the consequentiality question. Given the importance of consequentiality, additional research is needed on the placement of the consequentiality question along with other follow-up questions to assess whether the results from this study can be replicated and to explore alternative explanations for why the question order matters.

Turning to the econometrics, in naïve models without endogeneity controls, perceived consequentiality is found to be an important determinant of voting behaviour. However, using several different approaches to address endogeneity concerns, I find that consequentiality beliefs do not have a significant impact on voting. Together with the question ordering results, the results of this paper suggest that the new trend of including consequentiality follow-up questions in surveys may not be the panacea to SP validity issues.

A secondary aim of this paper is to provide the first estimates of the benefits of reducing BWAs. I find that households in Alberta have a positive WTP of $1.56 to $3.47 per year for 10 years to reduce a single BWA in small communities with less than 500 residents. However, households are not willing to pay to reduce BWAs in medium or large communities even though these BWA events affect substantially more people per rare event. Given that around 85% of respondents in the sample live in medium and large communities, the positive WTP for reducing BWA in small communities suggest a significant non-use benefit from reducing these BWAs in the province. Furthermore, the results suggest that people are relatively agnostic on the method used to improve drinking water reliability and do not have clear preferences for grey or green infrastructure. All these welfare estimates have been estimated using models that assume everyone has the same preferences and exploring preference heterogeneity remains an important aspect of future work.

One explanation for the counterintuitive finding on community sizes relates to the funding sources for drinking water infrastructure in Alberta. The provincial government is more involved in funding drinking water infrastructure projects in small communities compared to more populated communities.[22] Respondents may have internalized this funding breakdown and do not believe that provincial money should be spent on drinking water infrastructure in large communities.

Multiplying the range of estimated per household WTP values by the 1.6 million households in Alberta, the aggregate value of reducing one BWA in a small community is $2.5 to $5.5 million. The public benefits of eliminating all 50 BWAs experienced by small communities in a typical year is estimated to be $125 to $278 million. This research shows that although investing in drinking water infrastructure for small communities is expensive, even unaffected households are willing to shoulder part of the financial burden.

[22]The Alberta Municipal Water/Wastewater Partnership (AMWWP) is an example of drinking water funding programs that uses the population level of the community to determine funding levels. For communities under 1,000 people, the government funds 75% of a project and this percentage declines as the population of the community increases. Communities larger than 45,000 people are not eligible to receive money from this program.

Appendix 4.A Comparison of split samples

Table 4.A.1: Comparison summary of socio-demographic characteristics between question order split samples

Variable	Consequentiality Question Before (n = 375)	Consequentiality Question After (n = 382)	Difference	t-stat
Female	0.53	0.48	0.05	(1.35)
College	0.37	0.36	0.01	(0.19)
University	0.45	0.45	0.00	(-0.14)
High income	0.33	0.29	0.04	(1.27)
Income missing	0.11	0.12	-0.01	(-0.36)
Age index	1.00	1.03	-0.04	(-1.62)
Live in large community	0.54	0.43	0.11	(3.10)
Live in medium community	0.12	0.24	-0.11	(-4.09)

Appendix 4.B Model results with full sample and without weighting

Table 4.B.1: Probit model of perceived consequentiality without weights

	All Respondents	Consequentiality After Sample	Consequentiality Before Sample
BWA small	0.00206	0.00149	0.00111
	(0.00341)	(0.00454)	(0.00548)
BWA medium	-0.0822*	-0.160**	0.0175
	(0.0499)	(0.0676)	(0.0788)
BWA large	-0.0450	-0.236	0.257
	(0.110)	(0.149)	(0.177)
Forest management treatment	0.105	-0.0864	0.400**
	(0.109)	(0.150)	(0.175)
Tax amount	-0.000299	0.000275	-0.000829
	(0.000604)	(0.000805)	(0.000950)
Female	0.0914	0.109	0.0194
	(0.111)	(0.150)	(0.176)
College	-0.106	-0.148	-0.0928
	(0.160)	(0.216)	(0.250)
University	0.00150	0.189	-0.137
	(0.280)	(0.365)	(0.483)
High income	-0.196	-0.457	-0.0428
	(0.272)	(0.359)	(0.468)
Income missing	-0.338	-0.550	-0.283
	(0.272)	(0.360)	(0.467)
Age index	-0.398**	-0.0634	-0.881***
	(0.189)	(0.262)	(0.293)
Live in large community	0.160	0.189	0.157
	(0.114)	(0.156)	(0.176)
Consequentiality question after	-0.314***		
	(0.111)		
Partner information[a]	0.0377	0.272*	-0.238
	(0.110)	(0.148)	(0.175)
Constant	1.591***	1.118***	1.930***
	(0.318)	(0.428)	(0.483)
N	757	382	375
Log lik.	-341.4	-194.1	-136.8
Chi-squared	27.01	16.84	20.84

Notes: BWA = boil water advisory. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a] A dummy variable equal to 1 if the information treatment for the government agency partner information is presented at beginning of survey.

Table 4.B.2: Probit model of water management referendum voting without weights

	Base Model	Consequentiality Dummy Model	Consequential Model (C>1=1)	Inconsequential Model (C=0)
BWA small	0.00853***	0.00831***	0.00732**	0.0115
	(0.00293)	(0.00294)	(0.00325)	(0.00745)
BWA medium	-0.00310	0.00515	0.0155	-0.0504
	(0.0421)	(0.0423)	(0.0469)	(0.104)
BWA large	0.106	0.111	0.156	-0.256
	(0.0942)	(0.0945)	(0.104)	(0.239)
Forest management treatment	-0.0853	-0.0956	-0.0487	-0.421*
	(0.0938)	(0.0942)	(0.104)	(0.238)
Tax amount	-0.00242***	-0.00241***	-0.00252***	-0.00239*
	(0.000524)	(0.000526)	(0.000584)	(0.00131)
Female	0.358***	0.351***	0.363***	0.189
	(0.0955)	(0.0959)	(0.105)	(0.246)
College	0.235*	0.248*	0.223	0.427
	(0.134)	(0.135)	(0.146)	(0.378)
University	0.0679	0.0757	-0.113	1.378*
	(0.237)	(0.237)	(0.261)	(0.714)
High income	0.277	0.286	0.489*	-0.993
	(0.231)	(0.232)	(0.257)	(0.681)
Income missing	0.450*	0.479**	0.610**	-0.591
	(0.234)	(0.235)	(0.259)	(0.703)
Age index	0.0259	0.0603	0.114	-0.356
	(0.159)	(0.160)	(0.173)	(0.448)
Live in large community	0.0660	0.0493	0.0159	0.119
	(0.0958)	(0.0963)	(0.106)	(0.245)
Consequentiality Dummy (C>1=1)		0.361***		
		(0.125)		
Constant	-0.495**	-0.836***	-0.528*	-0.117
	(0.252)	(0.279)	(0.272)	(0.738)
Observations	757	757	622	135
Log likelihood	-498.9	-494.6	-408.9	-80.11
Chi-squared	51.10	59.55	44.07	17.75

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Table 4.B.3: Comparison of models that address potential endogeneity of consequentiality without weights

	Probit	Bivariate Probit	Control Function	Special Regressor[a]
BWA small	0.00831***	0.00788**	0.00842***	1.319***
	(0.00294)	(0.00316)	(0.00311)	(0.396)
BWA medium	0.00515	0.0153	0.00171	-0.201
	(0.0423)	(0.0474)	(0.0526)	(6.576)
BWA large	0.111	0.115	0.109	-1.093
	(0.0945)	(0.0942)	(0.101)	(12.50)
Forest management treatment	-0.0956	-0.104	-0.0916	-19.18
	(0.0942)	(0.0949)	(0.112)	(12.83)
Tax amount	-0.00241***	-0.00234***	-0.00242***	-
	(0.000526)	(0.000570)	(0.000553)	-
Female	0.351***	0.333***	0.355***	38.27***
	(0.0959)	(0.108)	(0.105)	(13.22)
College	0.248*	0.258*	0.244	31.64*
	(0.135)	(0.135)	(0.153)	(18.01)
University	0.0757	0.0792	0.0725	2.593
	(0.237)	(0.236)	(0.277)	(30.99)
High income	0.286	0.300	0.282	45.02
	(0.232)	(0.231)	(0.260)	(30.85)
Income missing	0.479**	0.509**	0.469*	63.15*
	(0.235)	(0.239)	(0.283)	(33.07)
Age index	0.0603	0.106	0.0429	17.62
	(0.160)	(0.187)	(0.209)	(26.43)
Live in large community	0.0493	0.0250	0.0568	-6.983
	(0.0963)	(0.110)	(0.125)	(14.94)
Consequentiality Dummy (C>1=1)	0.361***	0.808	0.461	61.94
	(0.125)	(0.952)	(0.779)	(157.2)
Residual			0.0550	
			(0.425)	
Constant	-0.836***	-1.250	-0.832***	-120.1
	(0.279)	(0.906)	(0.297)	(150.8)
ρ		-0.266		
		(0.603)		
N	757	757	757	757

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01. [a] The coefficient of the special regressor B is not shown here because it is normalized to one. Therefore, the other coefficients can be interpreted as WTP. The estimates are given by a two stage least squares regression linked to Step 4 of the special repressor approach outlined in the previous section.

Table 4.B.4: Willingness-to-pay to reduce boil water advisories without weights

	Probit Model (Full sample)	Bivariate Probit Model	Control Function Control Function	Special Regressor
BWA small	\$3.45**	\$3.36**	\$3.41**	\$1.32***
	(1.378)	(1.413)	(1.418)	(0.396)
BWA medium	\$2.14	\$6.53	\$6.89	-\$0.20
	(17.58)	(20.61)	(19.10)	(6.576)
BWA large	\$45.87	\$49.10	\$48.00	-\$1.09
	(39.78)	(41.52)	(39.50)	(12.50)
Consequentiality Dummy	\$149.80**	\$344.7	-\$14.02	\$61.94
(C>1=1)	(61.20)	(446.1)	(112.9)	(157.2)
N	757	757	757	757

Notes: BWA = boil water advisory. Standard errors in parentheses. * p <0.10, ** p <0.05, *** p <0.01.

Chapter 5

Conclusion

The research presented in this thesis provides many avenues for future investigations. The first paper focuses on intertemporal substitution of people participating in a recreational fishery. It would be useful to move beyond this specific resource and apply the dynamic structural model developed in this paper to other contexts where people respond to temporal variations in the quality of, and access to, recreational opportunities. Most recreation activities have a clear temporal component whether it be caused by natural variation in seasons or species abundance or government regulations restricting access during certain times of year. This extent of temporal change in recreation activities can be further disturbed by climate change and resource development projects. In Canada, climate change is already having impacts on ecological, social, and economic systems and behavioural models and valuation methods can play an important role in evaluating these impacts and helping inform appropriate adaptation.

Beyond intertemporal substitution, the second main focus of the first paper is on the appropriate value of time measure to use in travel cost models. The main finding of significant differences between the individual-specific value of time measure and the income-based measure, along with the accompanying robustness checks, raises important concerns about blindly following the 1/3 wage rule typically employed in the literature. This is a somber conclusion for travel cost practitioners who almost always include an income question in a RP survey or rely on linked census income metrics when individual information is not available. A useful research area moving forward is to assess for what type of individuals income provides meaningful information on how they value their time. For individuals whose income is not the most important determinant of the value of time, it is perhaps more appropriate to use a non-individual specific, average value of time (Fezzi et al., 2014).

Extending the methods implemented in the second paper to the field contexts may provide useful information on the external validity of the results. For public good contexts, the simple valuation task used in the experiment is a far cry from the more complex scenarios embedded in policy-relevant valuation questions. Would the reliability of WTA responses hold in a real-world setting when the question is consequential? Furthermore, respondents may not respond to questions using utility maximization decision rules which are at the heart of the theory of consequentiality. Gaining a better understanding if these types of nonconforming responses are more prevalent in the WTA context and implementing the relevant methods to assess these type of responses is an important area for future research. In the few cases where advisory referenda have been held in Canada, the WTP context is more common. For example, the most recent WTP example is the 2015 Metro Vancouver transit referendum where people voted no to an increase in the sales tax to pay for increased public transit. There are few cases where referenda on tax rebates have been held. In Canada, perhaps the closest we have come to a WTA referendum question is the *It's Your Future* survey conducted in 2004 in Alberta. Close to 290,000 Albertans provided their opinion on whether the windfall oil revenue for the government should be used to provide a refund to every Albertan or to make long-term investments in priority areas such as education and health care. The government did end up mailing out $400 cheques to each Albertan after the survey, although a direct link between the survey and the rebates was never explicitly made by the government. Further research on the consequentiality of WTA surveys in an environmental public good is warranted to generalize the findings of the experiment.

For private goods, the incentive compatibility literature suggests that we cannot collect unbiased preference information in most cases, regardless of whether a WTA or WTP question is posed (Carson and Groves, 2007). And yet the increased interest in PES programs reinforces the importance of private good settings for many WTA preference collection situations. The specific survey framing and follow-up questions used in the experiment is only the start of a broader effort required to assess the robustness of these approaches to influence and identify strategic behaviour. A clear limitation of these type of strategic behaviour follow-up questions is that they are largely inconsequential and rely on both the respondent acting strategically by misrepresenting their preferences in the valuation question as well as truthfully reporting this lie a few questions later. Refining the specific wording and format of these follow-up questions as well as broader approaches to handle level-k thinking remain important areas for future research.

The third paper contributes to the rapidly growing literature on the consequentiality of SP surveys. Mechanism design theory provides clear guidance on the importance of consequentiality in yielding truthful responses, and yet implementing the theory in applied work is not straightforward. The paper focuses on how to use responses to perceived consequentiality follow-up questions in econometric modeling of voting behaviour. Many of the same concerns with the strategic behaviour follow-up questions discussed in the second paper hold true for consequentiality follow-up questions. The results of the study provide additional evidence of the difficulty of eliciting meaningful consequentiality perceptions and appropriately modeling these responses. Given these difficulties, the use of consequentiality follow-up questions is no substitute for ensuring that the survey is consequential in the first place. Given the ongoing popularity of clearly inconsequential SP surveys, it is not clear how responses to consequentiality follow-up questions should be interpreted. Are misbeliefs about how policy makers may use the results of a clearly hypothetical study a valid measure of truthfullness?

Moving beyond *eliciting* consequentiality perceptions to *ensuring* survey consequentiality is an important task for SP practitioners going forward, recognizing that this is not possible in all survey contexts. Consequentiality is an important part of the reliability 'toolkit' to assess, and ultimately improve, SP surveys. Nevertheless, the broader hard and unglamorous work of good survey design including focus groups, cognitive interviews, and pre-tests, although largely unsung and perceived to be 'non-economic', continues to be the most important factor in improving the reliability of SP surveys.

Bibliography

Abbott, J. and Willard, D. (2017). Rights-Based Management for Recreational For-Hire Fisheries: Evidence from a Policy Trial. *Fisheries Research*, 196:106–116.

Abbott, J. K. and Fenichel, E. P. (2013). Anticipating adaptation: a mechanistic approach for linking policy and stock status to recreational angler behavior. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(8):1190–1208.

Adamowicz, W. L. (1994). Habit formation and variety seeking in a discrete choice model of recreation demand. *Journal of Agricultural and Resource Economics*, 19(1):19–31.

Aguiar, M. and Hurst, E. (2007). Life-Cycle Prices and Production. *American Economic Review*, 97(5):1533–1559.

Aguiar, M., Hurst, E., and Karabarbounis, L. (2013). Time Use during the Great Recession. *American Economic Review*, 103(5):1664–1696.

Alberta Health Service (2017). Inactive Health Advisories.

American Automobile Association (2015). 2015 Your Driving Costs. Technical report.

American Hotel & Lodging Association (2015). Lodging Industry Trends 2015. Technical report.

Arnott, R., de Palma, A., and Lindsey, R. (1993). A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand. *The American Economic Review*, 83(1):161–179.

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., and Shuman, H. (1993). Report of the NOAA Panel on Contingent Valuation. Technical report.

Ashenfelter, O. and Greenstone, M. (2004). Using Mandated Speed Limits to Measure the Value of a Statistical Life. *Journal of Political Economy*, 112(S1):S226–S267.

Association, A. A. (2016). 2016 Your Driving Costs. Technical report.

Baerenklau, K. A. and Provencher, B. (2005). Static modeling of dynamic recreation behavior: Implications for prediction and welfare estimation. *Journal of Environmental Economics and Management*, 50(3):617–636.

Baum, C. F. (2012). SSPECIALREG: Stata module to estimate binary choice model with discrete endogenous regressor via special regressor method.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, 75(299):493–517.

Bento, A. M., Goulder, L. H., Jacobsen, M. R., and Von Haefen, R. (2009). Distributional and Efficiency Impacts of Increased US Gasoline Taxes. *American Economic Review*, 99(3):667–699.

Bhat, C. and Pinjari, A. (2014). Multiple discrete-continuous choice models: a reflective analysis and a prospective view. In Hess, S. and Daly, A., editors, *Handbook of Choice Modelling*, chapter 19, pages 427–454. Edward Elgar Publishing.

Bhat, C. R. (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological*, 42(3):274–303.

Bishop, R. C. and Boyle, K. J. (2017). Reliability and Validity in Nonmarket Valuation. In *A Primer on Nonmarket Valuation*, The Economics of Non-Market Goods and Resources, pages 463–497. Springer, Dordrecht. DOI: 10.1007/978-94-007-7104-8_12.

Bishop, R. C., Boyle, K. J., Carson, R. T., Chapman, D., Hanemann, W. M., Kanninen, B., Kopp, R. J., Krosnick, J. A., List, J., Meade, N., Paterson, R., Presser, S., Smith, V. K., Tourangeau, R., Welsh, M., Wooldridge, J. M., DeBell, M., Donovan, C., Konopka, M., and Scherer, N. (2017). Putting a value on injuries to natural assets: The BP oil spill. *Science*, 356(6335):253–254.

Bockstael, N. E. and McConnell, K. E. (2007). *Environmental and Resource Valuation with Revealed Preferences: A Theoretical Guide to Empirical Models*. Springer Science & Business Media.

Bollino, C. A. (2009). The Willingness to Pay for Renewable Energy Sources: The Case of Italy with Socio-demographic Determinants. *The Energy Journal*, 30(2):81–96.

Bontemps, C. and Nauges, C. (2016). The Impact of Perceptions in Averting-decision Models: An Application of the Special Regressor Method to Drinking Water Choices. *American Journal of Agricultural Economics*, 98(1):297–313.

Boxall, P. C. and Englin, J. E. (2008). Fire and recreational values in fire-prone forests: Exploring an intertemporal amenity function using pooled RP-SP data. *Journal of Agricultural and Resource Economics*, 33(01):19–33.

Boyes, W. J. and Faith, R. L. (1993). Temporal Regulation and Intertemporal Substitution: The Effect of Banning Alcohol at College Football Games. *Public Choice*, 77(3):595–609.

Budziski, W. and Czajkowski, M. (2017). Hybrid Choice Models vs. endogeneity of indicator variables: a Monte Carlo investigation. In *International Choice Modelling Conference 2017*.

Bureau of Labor Statistics (2015). News Release USDL-14-1348. Technical report.

Burton, A. C., Carson, K. S., Chilton, S. M., and Hutchinson, W. G. (2007). Resolving questions about bias in real and hypothetical referenda. *Environmental and Resource Economics*, 38(4):513–525.

Bush, G., Hanley, N., Moro, M., and Rondeau, D. (2013). Measuring the local costs of conservation: a provision point mechanism for eliciting willingness to accept compensation. *Land Economics*, 89(3):490–513.

Calfee, J. and Winston, C. (1998). The value of automobile travel time: implications for congestion policy. *Journal of Public Economics*, 69(1):83–102.

Cameron, A. C. and Miller, D. L. (2015). A Practitioners Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Carlsson, F. (2010). Design of stated preference surveys: Is there more to learn from behavioral economics? *Environmental and Resource Economics*, 46(2):167–177.

Carson, R. T. and Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37(1):181–210.

Carson, R. T. and Groves, T. (2011). Incentive and information properties of preference questions: commentary and extensions. In *The International Handbook on Non-Market Environmental Valuation*. Edward Elgar Publishing.

Carson, R. T., Groves, T., and List, J. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, 1(1):171 – 207.

Carson, R. T., Hanemann, W. M., and Wegge, T. C. (2009). A nested logit model of recreational fishing demand in Alaska. *Marine Resource Economics*, 24(2):101–129.

Casey, J. F., Vukina, T., and Danielson, L. E. (1995). The Economic Value Of Hiking: Further Considerations Of Opportunity Cost Of Time In Recreational Demand Models. *Journal of Agricultural and Applied Economics*, 27(02).

Cash, S. (2015). The Facebook Suicides: Using Second-Price Auctions to Determine Attachment and Addiction to Facebook.

Castro, M., Bhat, C. R., Pendyala, R. M., and Jara-Daz, S. R. (2012). Accommodating multiple constraints in the multiple discrete-continuous extreme value (MDCEV) choice model. *Transportation Research Part B: Methodological*, 46(6):729–743.

Cesario, F. J. and Knetsch, J. L. (1970). Time Bias in Recreation Benefit Estimates. *Water Resources Research*, 6(3):700–704.

Chilton, S., Jones-Lee, M., McDonald, R., and Metcalf, H. (2011). Eliciting reliable willingness-to-accept responses.

Connolly, M. (2008). Here Comes the Rain Again: Weather and the Intertemporal Substitution of Leisure. *Journal of Labor Economics*, 26(1):73–100.

Corrigan, J. R., Kling, C. L., and Zhao, J. (2007). Willingness to pay and the cost of commitment: An empirical specification and test. *Environmental and Resource Economics*, 40(2):285–298.

Cummings, R. G., Elliott, S., Harrison, G. W., and Murphy, J. (1997). Are Hypothetical Referenda Incentive Compatible? *Journal of Political Economy*, 105(3):609–621.

Czajkowski, M., Vossler, C. A., Budziski, W., Winiewska, A., and Zawojska, E. (2015). Addressing empirical challenges related to the incentive compatibility of stated preference methods. Technical Report 2015-31, Faculty of Economic Sciences, University of Warsaw.

Dalya, A., De Jonga, G., Ibaneza, N., Batleya, R., and De Bok, M. (2008). Welfare measures from discrete choice models in the presence of income effect. *European Transport Conference 2008; Proceedings*.

Davis, L. W. (2008). The Effect of Driving Restrictions on Air Quality in Mexico City. *Journal of Political Economy*, 116(1):38–81.

DeSerpa, A. C. (1971). A Theory of the Economics of Time. *Economic Journal*, 81(324):828–46.

Doyon, M. and Bergeron, S. (2016). Understanding Strategic Behavior and Its Contribution to Hypothetical Bias When Eliciting Values for a Private Good. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 64(4):653–666.

Englin, J. and Cameron, T. A. (1996). Augmenting travel cost models with contingent behavior data. *Environmental and Resource Economics*, 7(2):133–147.

English, E., Leggett, C., and McConnell, K. (2015). Value of Travel Time and Income Imputation.

Espinosa-Goded, M., Barreiro-Hurl, J., and Ruto, E. (2010). What Do Farmers Want From Agri-Environmental Scheme Design? A Choice Experiment Approach. *Journal of Agricultural Economics*, 61(2):259–273.

Feather, P. and Shaw, W. D. (1999). Estimating the cost of leisure time for recreation demand models. *Journal of Environmental Economics and Management*, 38(1):49–65.

Fezzi, C., Bateman, I. J., and Ferrini, S. (2014). Using revealed preferences to estimate the Value of Travel Time to recreation sites. *Journal of Environmental Economics and Management*, 67(1):58–70.

Freeman III, A. M., Herriges, J. A., and Kling, C. L. (2014). *The Measurement of Environmental and Resource Values: Theory and Methods*. RFF Press, New York, New York, 3rd edition.

Glassdoor (2014). Q1 2014 Employment Confidence Survey. Technical report.

Graff Zivin, J. and Neidell, M. (2014). Temperature and the Allocation of Time: Implications for Climate Change. *Journal of Labor Economics*, 32(1):1–26.

Groothuis, P. A., Mohr, T. M., Whitehead, J. C., and Cockerill, K. (2017). Endogenous Consequentiality in Stated Preference Referendum Data: The Influence of the Randomly Assigned Tax Amount. *Land Economics*, 93(2):258–268.

Gulf of Mexico Fishery Management Council (2016). Reef Fish Recreational Management for Headboat Survey Vessels. Amendment 42 to the Fishery Management Plan for the Reef Fish Resources of the Gulf of Mexico. Technical report.

Haab, T., Huang, J., and Whitehead, J. (1999). Are hypothetical referenda incentive compatible? A comment. *Journal of Political Economy*, 107(1):186–196.

Haab, T. and McConnell, K. E. (2002). *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*. Edward Elgar Publishing.

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? *The American Economic Review*, 81(3):635–647.

Harrison, G. W. (2006). Making choice studies incentive compatible. In Kanninen, B. J., editor, *Valuing Environmental Amenities Using Stated Choice Studies*, number 8 in The Economics of Non-Market Goods and Resources, pages 67–110. Springer Netherlands.

Harrison, G. W. (2007). House money effects in public good experiments: Comment. *Experimental Economics*, 10(4):429–437.

Hartmann, W. R. (2006). Intertemporal effects of consumption and their implications for demand elasticity estimates. *Quantitative Marketing and Economics*, 4(4):325–349.

Hausman, J. (2012). Contingent Valuation: From Dubious to Hopeless. *Journal of Economic Perspectives*, 26(4):43–56.

Health Canada (2011). First Nations Communities with Drinking Water Advisories as of December 31, 2010.

Hensher, D., Rose, J., and Greene, W. (2015). *Applied Choice Analysis: A Primer*. Cambridge University Press, Cambridge, United Kingdom, 2nd edition edition.

Herriges, J., Kling, C., Liu, C.-C., and Tobias, J. (2010). What are the consequences of consequentiality? *Journal of Environmental Economics and Management*, 59(1):67–81.

Herriges, J. A. and Phaneuf, D. J. (2002). Inducing patterns of correlation and substitution in repeated logit models of recreation demand. *American Journal of Agricultural Economics*, 84(4):1076–1090.

Hess, S., Polak, J. W., Daly, A., and Hyman, G. (2007). Flexible substitution patterns in models of mode and time of day choice: new evidence from the UK and the Netherlands. *Transportation*, 34(2):213–238.

Homans, F. R. and Wilen, J. E. (1997). A Model of Regulated Open Access Resource Use. *Journal of Environmental Economics and Management*, 32(1):1–21.

Horne, P. (2006). Forest owners acceptance of incentive based policy instruments in forest biodiversity conservation a choice experiment based approach. *Silva Fennica*, 40(1):169–178.

Horowitz, J. K. and McConnell, K. E. (2002). A Review of WTA/WTP Studies. *Journal of Environmental Economics and Management*, 44(3):426–447.

Horowitz, J. L. (1992). A Smoothed Maximum Score Estimator for the Binary Response Model. *Econometrica*, 60(3):505–531.

Indigenous and Northern Affairs Canada (2017). Water in First Nation communities. Technical report, Indigenous and Northern Affairs Canada.

Industrial Economics Inc (2015). Air Travel Data Integration.

Interis, M. (2014). A challenge to three widely held ideas in environmental valuation. *Journal of Agricultural and Applied Economics*, 46(3):347–356.

Interis, M. and Petrolia, D. (2014). The effects of consequentiality in binary- and multinomial-choice surveys. *Journal of Agricultural and Resource Economics*, 39(2):201–216.

Jara-Diaz, S. R. and Videla, J. L. (1990). Welfare implications of the omission of income effect in mode choice models. *Journal of Transport Economics and Policy*, 24(1).

Johnston, R. J., Boyle, K. J., Adamowicz, W. V., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and Vossler, C. A. (2017). Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists*, 4(2):319–405.

Kaczan, D., Swallow, B. M., and Adamowicz, W. L. V. (2013). Designing a payments for ecosystem services (PES) program to reduce deforestation in Tanzania: An assessment of payment approaches. *Ecological Economics*, 95:20–30.

Kahle, D. and Wickham, H. (2013). ggmap: Spatial Visualization with ggplot2. *The R Journal*, 5(1):144–161.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–91.

Kalisa, T., Riddel, M., and Shaw, W. D. (2016). Willingness to pay to avoid arsenic-related risks: a special regressor approach. *Journal of Environmental Economics and Policy*, 5(2):143–162.

Kanjilal, M. (2015). *Costs of Wetland Restoration in the Canadian Prairies: Comparison of Simulated Opportunity Costs and farmers' WTA*. PhD thesis, University of Alberta; Department of Resource Economics and Environmental Sociology.

Karni, E. (1973). The Transactions Demand for Cash: Incorporation of the Value of Time into the Inventory Approach. *Journal of Political Economy*, 81(5):1216–1225.

Keeler, J. P., James, W. L., and Abdel-Ghany, M. (1985). The relative size of windfall income and the permanent income hypothesis. *Journal of Business & Economic Statistics*, 3(3):209–215.

Knetsch, J. L. (2007). Biased valuations, damage assessments, and policy choices: The choice of measure matters. *Ecological Economics*, 63(4):684–689.

Knetsch, J. L. (2010). Values of gains and losses: Reference states and choice of measure. *Environmental and Resource Economics*, 46(2):179–188.

Krishna, V. V., Drucker, A. G., Pascual, U., Raghu, P. T., and King, E. D. I. O. (2013). Estimating compensation payments for on-farm conservation of agricultural biodiversity in developing countries. *Ecological Economics*, 87:110–123.

Kuriyama, K. and Hanemann, W. M. (2006). The intertemporal substitution of recreation demand: A dynamic Kuhn-Tucker model with a corner solution. Unpublished Manuscript.

Laird, J. (2010). Transport welfare benefits in the presence of an income effect. In Hess, S. and Daly, A., editors, *Choice Modelling: The State-of-the-art and The State-of-practice*, chapter Chapter 18, pages 397–420. Emerald Group.

Landry, C. E., John, and List, A. (2007). Using ex ante approaches to obtain credible signals for value in contingent markets: Evidence from the field. *American Journal of Agricultural Economics*, pages 420–429.

Larson, D. M. and Shaikh, S. L. (2001). Empirical Specification Requirements for Two-Constraint Models of Recreation Choice. *American Journal of Agricultural Economics*, 83(2):428–440.

Latacz-Lohmann, U. and Schiizzi, S. (2005). Auctions for Conservation Contracts: A Review of the Theoretical and Empirical Literature (Project No: UKL/001/05). Report to the Scottish Executive Environment and Rural Affairs Department: Project No: UKL/001/05.

Latacz-Lohmann, U. and Van der Hamsvoort, C. (1997). Auctioning Conservation Contracts: A Theoretical Analysis and an Application. *American Journal of Agricultural Economics*, 79(2):407–418.

Leggett, C. G. (2015). Technical Report: Travel Cost Computation.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174.

Lew, D. K. and Larson, D. M. (2005). Accounting for stochastic shadow values of time in discrete-choice recreation demand models. *Journal of Environmental Economics and Management*, 50(2):341–361.

Lewbel, A. (2014). An Overview of the Special Regressor Method. In Racine, J. S., Su, L., and Ullah, A., editors, *The Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*. Oxford University Press. DOI: 10.1093/oxfordhb/9780199857944.013.002.

Lewbel, A., McFadden, D., and Linton, O. (2011). Estimating features of a distribution from binomial data. *Journal of Econometrics*, 162(2):170–188.

Lewbel, A. and Schennach, S. M. (2007). A simple ordered data estimator for inverse density weighted expectations. *Journal of Econometrics*, 136(1):189–211.

List, J. A. and Shogren, J. F. (2002). Calibration of willingness-to-accept. *Journal of Environmental Economics and Management*, 43(2):219–233.

Liu, C.-C., Herriges, J. A., Kling, C. L., Secchi, S., Nassauer, J. I., and Phaneuf, D. J. (2014). A comparison of value elicitation question formats in multiple-good contingent valuation. *Frontiers of Economics in China*, 9(1):85–108.

Lloyd-Smith, P. (2017). A New Approach to Calculating Welfare Measures in Kuhn-Tucker Demand Models.

Lloyd-Smith, P., Abbott, J., Adamowicz, W., and Willard, D. (2017). Decoupling the value of leisure time from labor market returns in travel cost models. Resource Economics and Environmental Sociology Working Paper, University of Alberta.

Lloyd-Smith, P. and Adamowicz, W. L. (2017). Can Stated Measures of Willingness-to-Accept be Valid? Evidence from Laboratory Experiments.

Lloyd-Smith, P., Schram, C., Adamowicz, W., and Dupont, D. (2016). Endogeneity of risk perceptions in averting behavior models. *Environmental and Resource Economics*.

Lupi, F. (2005). *Random Utility Models for Recreation Demand: Theory and Concepts for Welfare Measurement*. Chapter Notes for AEC 891.

Lusk, J. L., McLaughlin, L., and Jaeger, S. R. (2007). Strategy and Response to Purchase Intention Questions. *Marketing Letters*, 18(1/2):31–44.

Lusk, J. L. and Shogren, J. F. (2008). *Experimental Auctions: Methods and Applications in Economic and Marketing Research*. Cambridge University Press.

McFadden, D. and Train, K. (2017). *Contingent Valuation of Environmental Goods: A Comprehensive Critique*. Edward Elgar Publishing, Cheltenham, UK.

McFadden, D. L. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*. Wiley, New York.

Mendelsohn, R. and Neumann, J. E. (1999). The Impact of Climate Change on Outdoor Recreation. In *The Impact of Climate Change on the United States Economy*, pages 267–288. Cambridge University Press. Google-Books-ID: 8NMUQ3h5i9IC.

Messer, K. D., Poe, G. L., Rondeau, D., Schulze, W. D., and Vossler, C. A. (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. *Journal of Public Economics*, 94(34):308–317.

Mitchell, R. and Carson, R. T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. RFF Press.

Mitchell, R. C. (2002). On designing constructed markets in valuation surveys. *Environmental and Resource Economics*, 22(1-2):297–321.

Mler, K. (1974). *Environmental Economics: A Theoretical Inquiry*. Johns Hopkins University Press for Resources for the Future.

Moeltner, K. and Englin, J. (2004). Choice behavior under time-variant quality. *Journal of Business & Economic Statistics*, 22(2):214–224.

Mulligan, C. B. (1997). Scale Economies, the Value of Time, and the Demand for Money: Longitudinal Evidence from Firms. *Journal of Political Economy*, 105(5):1061–1079.

Oehlmann, M. and Meyerhoff, J. (2017). Stated preferences towards renewable energy alternatives in Germany do the consequentiality of the survey and trust in institutions matter? *Journal of Environmental Economics and Policy*, 6(1):1–16.

Ovaskainen, V., Neuvonen, M., and Pouta, E. (2012). Modelling recreation demand with respondent-reported driving cost and stated cost of travel time: A Finnish case. *Journal of Forest Economics*, 18(4):303–317.

Palmquist, R. B., Phaneuf, D. J., and Smith, V. K. (2010). Short run constraints and the increasing marginal value of time in recreation. *Environmental and Resource Economics*, 46(1):19–41.

Parsons, G. R. (2003). The Travel Cost Model. In Champ, P. A., Boyle, K. J., and Brown, T. C., editors, *A Primer on Nonmarket Valuation*, number 3 in The Economics of Non-Market Goods and Resources, pages 269–329. Springer Netherlands.

Phaneuf, D. J. and Herriges, J. A. (1999). Choice set definition issues in a Kuhn-Tucker model of recreation demand. *Marine Resource Economics*, 14(4):343–355.

Phaneuf, D. J., Kling, C. L., and Herriges, J. A. (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand. *Review of Economics and Statistics*, 82(1):83–92.

Phaneuf, D. J. and Smith, V. K. (2005). Chapter 15 Recreation Demand Models. In Vincent, K.-G. M. a. J. R., editor, *Handbook of Environmental Economics*, volume 2 of *Valuing Environmental Changes*, pages 671–761. Elsevier.

Pinjari, A. R. and Bhat, C. R. (2011). Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: Application to residential energy consumption analysis.

Plott, C. R. and Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *The American Economic Review*, 95(3):530–545.

Porras, I. and Hope, R. (2005). Using stated choice methods in the design of payments for environmental services schemes.

Provencher, B. and Bishop, R. C. (1997). An estimable dynamic model of recreation behavior with an application to Great Lakes angling. *Journal of Environmental Economics and Management*, 33(2):107–127.

Ray, R., Sanes, M., and Schmitt, J. (2013). No-Vacation Nation Revisited. Technical report, Center for Economic and Policy Research.

Riddel, M. (2011). Uncertainty and measurement error in welfare models for risk changes. *Journal of Environmental Economics and Management*, 61(3):341–354.

Sayman, S. and Onculer, A. (2005). Effects of study design characteristics on the WTAWTP disparity: A meta analytical framework. *Journal of Economic Psychology*, 26(2):289–312.

Shaikh, S. L. and Larson, D. M. (2003). A Two-Constraint Almost Ideal Demand Model of Recreation and Donations. *The Review of Economics and Statistics*, 85(4):953–961.

Shaw, W. D. and Loomis, J. B. (2008). Frameworks for analyzing the economic effects of climate change on outdoor recreation. *Climate Research*, 36(3):259–269.

Shi, J. and Skuterud, M. (2015). Gone Fishing! Reported Sickness Absenteeism and the Weather. *Economic Inquiry*, 53(1):388–405.

Small, K. A. and Verhoef, E. (2007). *The Economics of Urban Transportation*. Routledge, New York, New York, 2 edition.

Smith, V. K. and Banzhaf, H. S. (2004). A Diagrammatic Exposition of Weak Complementarity and the Willig Condition. *American Journal of Agricultural Economics*, 86(2):455–466.

Smith, V. K., Desvousges, W. H., and McGivney, M. P. (1983). The opportunity cost of travel time in recreation demand models. *Land Economics*, 59(3):259–278.

Southgate, D., Haab, T., Lundine, J., and Rodrguez, F. (2010). Payments for environmental services and rural livelihood strategies in Ecuador and Guatemala. *Environment and Development Economics*, 15(01):21–37.

Statistics Canada (2011). Focus on Geography Series.

Swait, J., Adamowicz, W., and Bueren, M. v. (2004). Choice and temporal welfare impacts: incorporating history into discrete choice models. *Journal of Environmental Economics and Management*, 47(1):94–116.

Thaler, R. H. and Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, 36(6):643–660.

Tseng, Y.-Y. and Verhoef, E. (2008). Value of time by time of day: A stated-preference study. *Transportation Research Part B: Methodological*, 42(7-8):607–618.

Tuncel, T. and Hammitt, J. K. (2014). A new meta-analysis on the WTP/WTA disparity. *Journal of Environmental Economics and Management*, 68(1):175–187.

U.S. Department of the Interior (2016). Deepwater Horizon Response and Restoration, Administrative Record. Technical report.

Van Nostrand, C., Sivaraman, V., and Pinjari, A. R. (2013). Analysis of long-distance vacation travel demand in the United States: a multiple discretecontinuous choice framework. *Transportation*, 40(1):151–171.

Vasquez Lavin, F. and Hanemann, W. M. (2008). Functional forms in discrete/continuous choice models with general corner solution. CUDARE Working Paper Series 1078, University of California at Berkeley, Department of Agricultural and Resource Economics and Policy.

Villanueva, A. J., Glenk, K., and Rodrguez-Entrena, M. (2017). Protest Responses and Willingness to Accept: Ecosystem Services Providers Preferences towards Incentive-Based Schemes. *Journal of Agricultural Economics*, pages n/a–n/a.

von Haefen, R. H. (2007). Empirical strategies for incorporating weak complementarity into consumer demand models. *Journal of Environmental Economics and Management*, 54(1):15 – 31.

von Haefen, R. H. and Phaneuf, D. J. (2005). Kuhn-Tucker Demand System Approaches to Non-Market Valuation. In Scarpa, R. and Alberini, A., editors, *Applications of Simulation Methods in Environmental and Resource Economics*, pages 135–157. Springer Netherlands, Dordrecht.

von Haefen, R. H., Phaneuf, D. J., and Parsons, G. R. (2004). Estimation and Welfare Analysis with Large Demand Systems. *Journal of Business & Economic Statistics*, 22(2):194–205.

Vossler, C. A., Doyon, M., and Rondeau, D. (2012). Truth in consequentiality: Theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, 4(4):145–171.

Vossler, C. A. and Watson, S. B. (2013). Understanding the consequences of consequentiality: Testing the validity of stated preferences in the field. *Journal of Economic Behavior & Organization*, 86:137–147.

Wales, T. J. and Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, 21(3):263–285.

Wang, H. and He, J. (2011). Estimating individual valuation distributions with multiple bounded discrete choice data. *Applied Economics*, 43(21):2641–2656.

Wang, H. and Whittington, D. (2005). Measuring individuals' valuation distributions using a stochastic payment card approach. *Ecological Economics*, 55(2):143–154.

Whittington, D., Adamowicz, W., and Lloyd-Smith, P. (2017). Asking Willingness-to-Accept Questions in Stated Preference Surveys: A Review and Research Agenda. *Annual Review of Resource Economics*, 9(1):317–336.

Whittington, D. and Pagiola, S. (2012). Using contingent valuation in the design of payments for environmental services mechanisms: A review and assessment. *World Bank Research Observer*, 27(2):261–287.

Willig, R. D. (1976a). Consumer's surplus without apology. *The American Economic Review*, 66(4):589–597.

Willig, R. D. (1976b). Consumer's surplus without apology. *The American Economic Review*, 66(4):589–597.

Wolff, H. (2014). Value of time: Speeding behavior and gasoline prices. *Journal of Environmental Economics and Management*, 67(1):71–88.

Wooldridge, J. M. (2015). Control Function Methods in Applied Econometrics. *Journal of Human Resources*, 50(2):420–445.

Supplementary Material A: Two Page Initial Headboat Fishery Survey

GULF OF MEXICO HEADBOAT ANGLER SURVEY

Please take 3 to 5 minutes to answer the following questions. This survey is being conducted to understand angler experiences in the Gulf of Mexico. Your participation is completely voluntary. If you prefer not to answer a question, feel free to skip it and go on to the next question. The information you provide will only be used for research purposes. All responses are confidential. No one will be identified in any reports coming out of the survey.

This survey research is being conducted by Joshua Abbott (Arizona State University). If you have any questions about this research, please contact me at joshua.k.abbott@asu.edu. If you have any questions about your rights as a participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Institutional Review Board, through the ASU Office of Research Integrity and Assurance, at (480) 965-6788.

By continuing to the survey, I acknowledge that I am at least 18 years old, have read the above information, and provide my consent to participate under the terms above.

---

1. Today's date: month________/day________/year__________

2. Current time: _____:______ am/pm

3. Name of the company/boat: ___________________________ /_________________________

4. Duration of fishing trip (to nearest hour): _____________ hours

5. Counting yourself, how many people were in your personal fishing group today?

☐☐ people

6. How many years have you been saltwater fishing? ☐☐ years

7. How often do you go offshore saltwater fishing in a typical year?
□ Less than once □ 1-2 times □ 3-6 times □ More than 6 times a year

8. Do you own a boat that is capable of fishing in the Gulf of Mexico (including nearshore fishing)?
□ Yes □ No

9. How many nights away from home (if any) are you spending on this trip? ☐☐ nights

10. Are you a repeat customer of this company? □ Yes □ No □ Don't know

11. What is the primary purpose of your trip?
□ This fishing trip □ Beach recreation □ Visiting family □ Other (specify)________________

Please turn over (more on back) →

12. How far in advance did you book today's fishing trip?
□ A week or less □ More than a week but less than a month □ More than a month

13. How much did you pay per person in fees and tips for today's trip?
Fee__________________ Tips____________________

14. Did you pay an additional surcharge for red snapper or gag grouper that you landed (Y/N)? If so, how much did you pay per fish? ___________________________

15. By species (if possible) how many fish did you personally catch and keep on today's trip?

Species	# caught	# kept
Red snapper		
Gag grouper		
All other catch		

16. How satisfied are you with today's fishing experience?
□ Very satisfied □ Satisfied □ Neutral □ Dissatisfied □Very dissatisfied

Comments____

____

17. What is your gender? □ Male □ Female

18. What year were you born? 1 9 [ ][ ]

19. What is the 5 digit US zip code (or country) of your place of residence? [ ][ ][ ][ ][ ]

20. What was your household's total income before taxes in 2013?
□ Less than \$24,999 □ \$75,000 - \$99,999 □ \$150,000 - \$199,999
□ \$25,000 - \$49,999 □ \$100,000 - \$124,999 □ \$200,000 - \$249,999
□ \$50,000 - \$74,999 □ \$125,000 - \$149,999 □ Greater than \$250,000

Thank you for your participation! If possible, we would like to contact you via email to complete a brief internet questionnaire about your opinions on today's fishing experience. Please provide an email below where we can contact you. Your email will be kept confidential, will not be given away or sold to anyone, and will not be used for marketing by any company. *By providing an email, you are providing your consent for us to contact you for this follow-up internet survey.*

Email (please print): ____________________________________@_________________________

<u>By providing your email, you will be entered into a drawing for a free fishing trip!</u>

Supplementary Material B: Main Headboat Recreational Fishing Survey (2015 version)

Introduction Screen

Gulf of Mexico Recreational Angler Headboat Survey

Study Overview

Offshore Saltwater Fishing in the Gulf of Mexico

Who is administering this survey?
Joshua Abbott, Associate Professor at Arizona State University, Email: Joshua.k.abbott@asu.edu.
Pat Lloyd-Smith, Graduate Student at University of Alberta, Email: lloydsmi@ualberta.ca.
Vic Adamowicz, Professor at the University of Alberta, Email: vic.adamowicz@ualberta.ca

Why are you contacting me? We are contacting you because you or someone sharing your email address completed a brief questionnaire we administered during a fishing trip in the Gulf of Mexico in 2015. By providing your email on this questionnaire, you authorized us to contact you for this follow-up survey.

What is the purpose of the survey? The goal of this survey is to better understand the opinions on recreational fishery management of people who took fishing trips in the Gulf of Mexico. This information could be used to create more valuable recreational fishing opportunities through changes in management policies in the Gulf of Mexico.

How much time will this take? The survey should take about 25 minutes of your time.

What are the benefits or risks to you? Participants in this survey will be entered into a drawing for a $100 Amazon gift card. Survey participants will assist the researchers in obtaining estimates of the public's views on fishery management policies and how these policies affect recreational fishing trips. There are no known or anticipated risks associated with participation in this study.

Confidentiality: All information you provide is considered confidential and grouped with responses from other participants. Email addresses will not be associated with survey responses. Access to the data will be restricted to Dr. Abbott and the investigators working on his research team.

Publication of Results: Grouped results of this study may be published in professional journals and presented at conferences. Feedback about this study will be available December 2016 from the principal investigator using the contact information provided below.

Withdrawal: Participation in this study is voluntary. If you wish, you may decline to answer any questions or participate in any component of the study. Further, you may decide to withdraw from this study at any time by declining to answer questions. Once the survey has been completed you cannot withdraw the information you provided.

Further information: For questions about this survey research contact Dr. Joshua Abbott (Arizona State University) at Joshua.k.abbott@asu.edu. If you have questions about your rights as a participant in this research or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Institutional Review Board through the ASU Office of Research Integrity and Assurance at (480) 965-6788. The plan for this study has also been reviewed for its adherence to ethical guidelines by a Research Ethics Board at the University of Alberta.

By continuing to the survey, I acknowledge that I am at least 18 years old, have read the above information, and provide my consent to participate under the terms above.

Browser Meta Info
This question will not be displayed to the recipient.
Browser: Chrome
Version: 58.0.3029.110
Operating System: Windows NT 10.0
Screen Resolution: 1920x1080
Flash Version: -1
Java Support: 0
User Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari/537.36

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

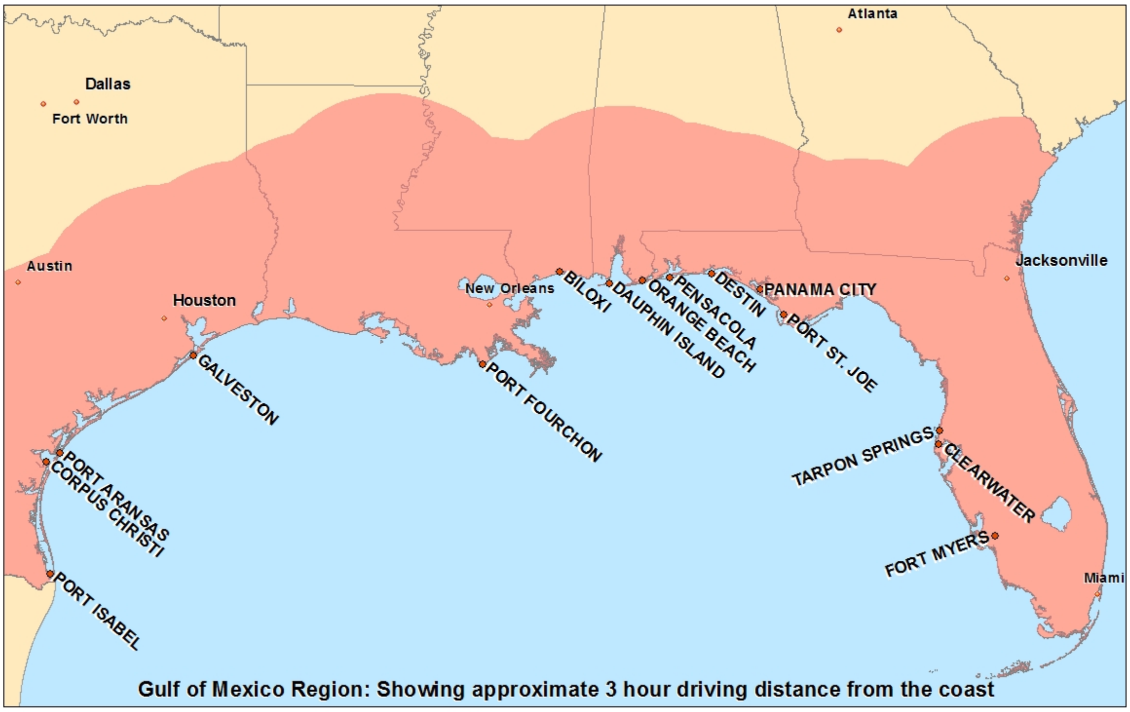
Study Overview

Are you the individual that completed the short survey onboard the headboat trip on ${e://Field/SurveyDate}?

- Yes
- No

What is the 5 digit US zip code, Canadian postal code, or country of your permanent place of residence?

For the rest of the survey, please consider the Gulf of Mexico Region to include the area within a 3 hour drive to the Gulf of Mexico in the five US states of Texas, Louisiana, Mississippi, Alabama, and Florida.



How often are you in the Gulf of Mexico Region?

- I live here year round
- I have a vacation/secondary home in the region and spend part of the year in the region.
- I do not have a vacation/secondary home, but I regularly visit at least once a year
- I visit the region less than once a year

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

What is the address of your vacation/secondary home?

- Texas
- Louisiana
- Mississippi
- Alabama
- Florida

City/Town	
ZIP Code	

What time of year are you typically in the Gulf of Mexico Region? Select all that apply.

- January to end of May
- June
- July to end of August
- September to end of December
- Time of year varies year to year

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

What is the best description of your occupational status? Please select only one response.

- Employed full-time
- Employed part-time
- Self-employed
- Not currently working
- Fully retired

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

We would now like to ask you about when you spent your vacation time in 2015.

A vacation day is any day spent not working that is not your usual days off, such as weekends or holidays (i.e. Independence Day, Labor Day, etc.).

For the following four time periods, please fill in the total number of vacation days you took in 2015 as well as the total number of nights you stayed away from your primary residence on these vacation days.

	January to end of May Holidays: New Year's, Spring break, Easter, Memorial day	June	July to end of August Holidays: Independence Day	September to end of December Holidays: Labor Day, Columbus Day, Thanksgiving, Christmas
Total number of vacation days				
Total number of nights away on trips from permanent home or secondary residence				

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

How many paid days off a year do you receive? Please enter zero if you do not receive any paid days off.

Paid days off

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Can you ‘bank’ these paid days off between years?

- Yes
- No

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds

Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Recreation

Recreation

For this survey “offshore fishing” refers to fishing in federal ocean waters (beyond 9 miles in FL and TX and 3 miles in other Gulf States.)

Besides offshore fishing, what other leisure activities have you done in the Gulf of Mexico Region? Check all that apply.

- [ ] Nearshore fishing (i.e. pier, jetty, or small craft within 3 miles of shore)
- [ ] Freshwater fishing
- [ ] Hunting
- [ ] Golf
- [ ] Hiking/Camping
- [ ] Beach recreation
- [ ] Water recreation (e.g., kayaking, scuba diving)
- [ ] Visit cultural/tourist sites
- [ ] Other activities (please specify) ____

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

How many years have you been offshore fishing in the Gulf of Mexico Region?

- ( ) Years ____
- ( ) I've only fished in the Gulf of Mexico once

What time(s) of year do you never consider going offshore fishing in the Gulf of Mexico? Check all that apply.

- [ ] January to end of May
- [ ] June
- [ ] July to end of August
- [ ] September to end of December

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Besides offshore fishing in the Gulf of Mexico, where else do you go fishing?

- ☐ I only fish in the Gulf of Mexico
- ☐ Saltwater fishing in the Atlantic Ocean
- ☐ Saltwater fishing in the Pacific Ocean
- ☐ Freshwater fishing in North America
- ☐ Fishing outside of North America

What species of fish are you interested in catching in the Gulf of Mexico? Select all that apply.

- ☐ Red snapper
- ☐ Gag grouper
- ☐ Dolphinfish (e.g., dorado, mahi-mahi)
- ☐ King Mackerel
- ☐ Triggerfish
- ☐ Redfish
- ☐ Amberjack
- ☐ Speckled trout
- ☐ Tuna
- ☐ Cobia
- ☐ Billfish (e.g., marlin)
- ☐ Shark
- ☐ Vermillion snapper
- ☐ Other species ______
- ☐ I'm not strongly motivated to catch any particular species
- ☐ I don't know enough about Gulf of Mexico species to have an opinion

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Regarding all the species you selected, which species are you most interested in catching?

- ○ » Red snapper
- ○ » Gag grouper
- ○ » Dolphinfish (e.g., dorado, mahi-mahi)
- ○ » King Mackerel
- ○ » Triggerfish
- ○ » Redfish
- ○ » Amberjack
- ○ » Speckled trout
- ○ » Tuna
- ○ » Cobia
- ○ » Billfish (e.g., marlin)
- ○ » Shark
- ○ » Vermillion snapper
- ○ » Other species ______
- ○

» I’m not strongly motivated to catch any particular species

» I don’t know enough about Gulf of Mexico species to have an opinion

Headboat Fishing Trips

We are now going to ask you some questions about headboat trips. A headboat is a vessel that takes groups of more than 6 passengers on recreational fishing trips in federal waters of the US Gulf of Mexico. Headboats typically charge an individual fee per customer for their trips.

Please indicate whether you agree or disagree with the following statements

	Strongly Agree	Somewhat Agree	Neither Agree nor Disagree	Somewhat Disagree	Strongly Disagree
I have changed when I normally go fishing because of season length restrictions for my preferred species.					
If I could retain more fish of my preferred species on each trip, I would increase my headboat fishing trips in the Gulf of Mexico					
If I could retain a greater variety of the fish I catch, I would increase my headboat fishing in the Gulf of Mexico					
I usually eat the fish I catch					

Strongly Agree Somewhat Agree Neither Agree nor Disagree Somewhat Disagree Strongly Disagree

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

How important were the following factors in your decision of whether <u>and</u> when to go headboat fishing. Please use the following scale from unimportant to very important.

Please select one response for each item

	Unimportant	Of Little Importance	Moderately Important	Important	Very Important

Holidays (i.e. Memorial Day, Independence Day, Thanksgiving)	○	○	○	○	○
Already traveling in the region for another purpose	○	○	○	○	○
The ability to retain catch of preferred species	○	○	○	○	○
Number of other anglers on the boat (i.e. crowdedness)	○	○	○	○	○
Ability to take time away from work or other commitments	○	○	○	○	○
Overall trip cost	○	○	○	○	○

Thinking of a typical headboat fishing trip, how did you decide which specific headboat company to take?

- ○ Past experience with the company
- ○ Tourism magazine/brochure
- ○ Websites/search engines
- ○ Word of mouth
- ○ Friends/family chose the trip

These page timer metrics will not be displayed to the recipient.

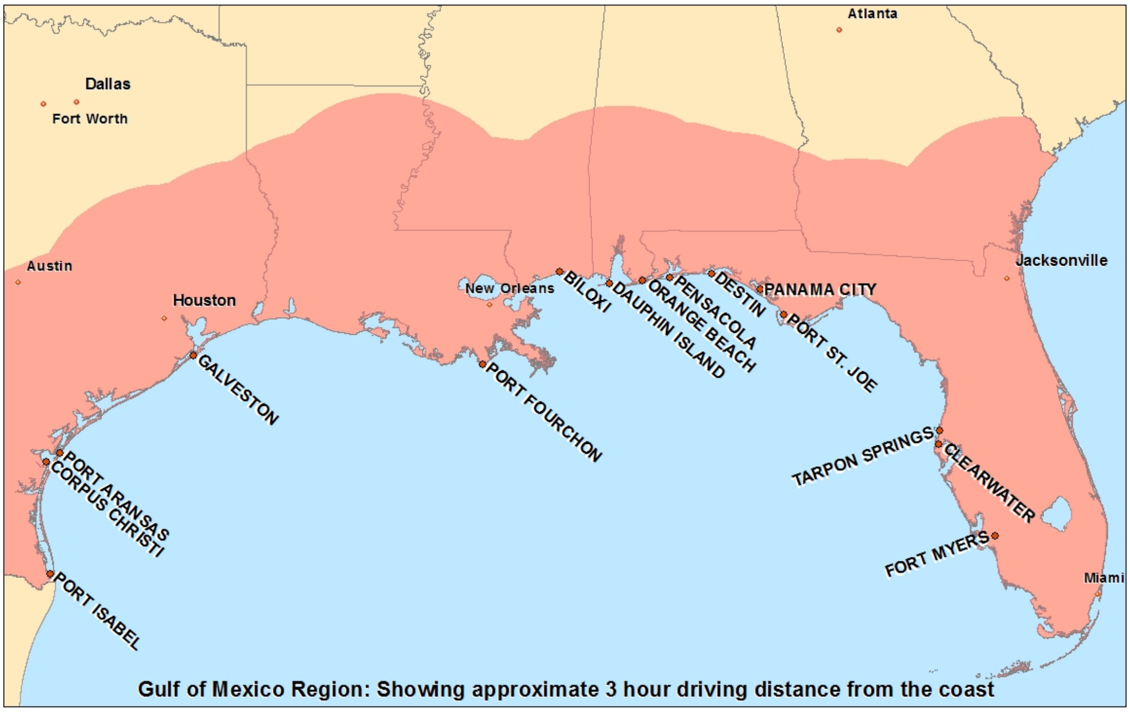
First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

In what locations did you go headboat fishing in 2015? Check all that apply. A map is provided below for reference.



Florida

- ☐ Clearwater/St. Pete Beach/Madeira Beach
- ☐ Destin
- ☐ Fort Myers Beach
- ☐ Panama City
- ☐ Pensacola/Gulf Breeze
- ☐ Port St Joe
- ☐ Tarpon Springs

Alabama

- ☐ Dauphin Island
- ☐ Orange Beach

Lousiana/Mississippi

- ☐ Biloxi
- ☐ Port Fourchon/Golden Meadow

Texas

- ☐ Galveston
- ☐ South Padre Island/Port Isabel
- ☐ Port Aransas

☐ Other (please enter location below)

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Trip Recall

We are now going to ask you for information about the headboat fishing trips you took to the Gulf of Mexico in 2015 in different time frames.

In responding to the following questions, please consider a trip as leaving your permanent residence or secondary/vacation home in order to go headboat fishing and then returning to that residence or secondary home some time later.

Please fill in the following table with the number of headboat trips and trip details for each of the four time periods. Enter '0' if a certain time period does not apply.

	January to end of May Holidays: New Year's, Spring break, Easter, Memorial day	June	July to end of August Holidays: Independence Day	September to end of December Holidays: Labor Day, Columbus Day, Thanksgiving, Christmas
Total number of headboat trips				
Total number of nights away on trips from permanent home or secondary residence				

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Detailed Recall A

Below is a summary of the total number of headboat trips you took in the Gulf of Mexico in 2015. Is this information correct? Please make any corrections and/or click next to continue. For the rest of the survey, all blank responses will be interpreted as zeros.

	January to end of May	June	July to end of August	September to end of December
Total number of headboat trips	${e://F	${e://F	${e://F	${e://F

Detailed Recall

Total Number of Headboat trips in 2015

	January to May	June	July to end of August	September to end of December

Total number of headboat trips	${e://Field/trips1}	${e://Field/trips2}	${e://Field/trips3}	${e://Field/trips4}

For each time period you took headboat trips, please provide a breakdown of the lengths of trips you took. Please make sure the breakdown of trips is equal to the total number of trips provided above.

For the ${e://Field/trips1} headboat trips you took in January to May 2015, please provide a breakdown of the lengths of trips you took.

Number of partial day trips (4 to 8 hrs)	0
Number of full day trips (8 to 15 hrs)	0
Number of more than one day trips (>15 hrs)	0
Total	0

For the ${e://Field/trips2} headboat trips you took in June 2015, please provide a breakdown of the lengths of trips you took.

Number of partial day trips (4 to 8 hrs)	0
Number of full day trips (8 to 15 hrs)	0
Number of more than one day trips (>15 hrs)	0
Total	0

For the ${e://Field/trips3} headboat trips you took in July and August, 2015, please provide a breakdown of the lengths of trips you took.

Number of partial day trips (4 to 8 hrs)	0
Number of full day trips (8 to 15 hrs)	0
Number of more than one day trips (>15 hrs)	0
Total	0

For the ${e://Field/trips4} headboat trips you took in September to December, 2015, please provide a breakdown of the lengths of trips you took.

Number of partial day trips (4 to 8 hrs)	0
Number of full day trips (8 to 15 hrs)	0
Number of more than one day trips (>15 hrs)	0
Total	0

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

For each time period you took headboat trips, please indicate the fish species retained. Click on the fish image to enlarge.

	January to end of May	June	July to end of August	September to end of December
Red Snapper 	☐	☐	☐	☐
Gag Grouper 	☐	☐	☐	☐
Other	☐	☐	☐	☐
Don't remember	☐	☐	☐	☐

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

When did you take the headboat trips from the different locations you visited in 2015? Select all that apply for each time period.

	» January to end of May	» June	» July to end of August	» September to end of December
» Clearwater/St. Pete Beach/Madeira Beach	☐	☐	☐	☐
» Destin	☐	☐	☐	☐
» Fort Myers Beach	☐	☐	☐	☐
» Panama City	☐	☐	☐	☐
» Pensacola/Gulf Breeze	☐	☐	☐	☐
» Port St Joe	☐	☐	☐	☐
» Tarpon Springs	☐	☐	☐	☐
» Dauphin Island	☐	☐	☐	☐
» Orange Beach	☐	☐	☐	☐

	» January to end of May	» June	» July to end of August	» September to end of December
» Biloxi				
» Port Fourchon/Golden Meadow				
» Galveston				
» South Padre Island/Port Isabel				
» Port Aransas				
» Other (please enter location below)				

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

What would you have done questions

If you had not been able to retain red snapper on your January to end of May trips what do you think you would have done?

	Postpone the headboat trip to a later date when red snapper fishing was allowed	Stay at home	Take a fishing trip at an alternative location	Spend the time and money on an alternative leisure activities in the area	Still take the headboat trip and fish for other species
January to end of May					

If you had not been able to retain gag grouper on your January to end of May trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
January to end of May					

If you had not been able to retain red snapper and gag grouper on your January to end of May trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
January to end of May					

If you had not been able to retain red snapper on your June trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
June	○	○	○	○	○

If you had not been able to retain gag grouper on your June trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
June	○	○	○	○	○

If you had not been able to retain red snapper and gag grouper on your June trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
June	○	○	○	○	○

If you had not been able to retain red snapper on your July to end of August trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
July to end of August	○	○	○	○	○

If you had not been able to retain gag grouper on your July to end of August trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
July to end of August	○	○	○	○	○

If you had not been able to retain red snapper and gag grouper on your July to end of August trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
July to end of August	○	○	○	○	○

If you had not been able to retain red snapper on your September to end of December trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
September to end of December	○	○	○	○	○

If you had not been able to retain gag grouper on your September to end of December trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
September to end of December	○	○	○	○	○

If you had not been able to retain red snapper and gag grouper on your September to end of December trips what do you think you would have done?

	» Still take the headboat trip and fish for other species	» Postpone the headboat trip to a later date when red snapper fishing was allowed	» Take a fishing trip at an alternative location	» Spend the time and money on an alternative leisure activities in the area	» Stay at home
September to end of December	○	○	○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Recall Headboat Trips Part 2

At the beginning of 2015, what had you planned regarding the number of headboat fishing trips for the year and roughly when you would take them? For the timing of the trips, please think in terms of the four time periods in the table below

Time Period	January to end of May	June	July to end of August	September to end of December

- I had planned the number of my fishing trips but I did not know what season(s) I would go
- I had planned the seasonal timing of my fishing trips but I did not know how many trips I would take
- I had planned the number and seasonal timing of my fishing trips
- I did not know the number or seasonal timing my fishing trips

Including yourself, what was the total number of people in your personal fishing trip party (not including crew or other customers on the boat) on a typical headboat fishing trip in 2015?

Number of people

Thinking of a typical trip during which you went headboat fishing in 2015, was fishing the

- Sole purpose of your trip
- Major purpose of your trip
- One of many equally important reasons
- An incidental reason or spur of the moment decision

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

What were some other purposes of the trip? Check all that apply.

- Only fishing
- Beach recreation
- Visit cultural/tourism sites
- Visit family or friends
- Business/work
- Other (please specify)

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Thinking about headboat trips in the Gulf of Mexico in the five years before 2015, when have you taken headboat trips in previous years? Select all that apply.

- January to end of May
- June
- July to end of August
- September to end of December
- I did not take any headboat trips in the five years before 2015

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

When did you take the majority of the trips in the five years before 2015?

- » January to end of May
- » June
- » July to end of August
- » September to end of December
- » I did not take any headboat trips in the five years before 2015

People have many reasons for why they do not go headboat fishing more often. For yourself, how important are the following reasons for not taking additional headboat trips in the Gulf of Mexico? Please use the following scale from unimportant to very important.

Please select one response for each item

	Unimportant	Of Little Importance	Moderately Important	Important	Very Important
I'd rather do something else					
Limits on retention of fish					
The money it costs to go fishing					
The time it takes to go fishing					
The quality of the fishing experience					

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Before this survey, were you aware of the following details of the red snapper and gag grouper fishery in the Gulf of Mexico?

	Yes, I was aware	No, I was not aware

There is a two fish bag limit on the number of red snapper or gag grouper a recreational angler can retain	○	○
Anglers can only land red snapper or gag grouper in federal waters during the seasons set by federal fisheries managers	○	○
There is a pilot program to allow headboat operators to fish for red snapper and gag grouper outside of the official recreational seasons set by federal fisheries managers	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

CB-CE Intro Gag Grouper

Managing the Gulf of Mexico Fisheries

Fisheries managers are considering changes to recreational fishing policy for headboats in the Gulf of Mexico that may impact ${e://Field/Fish} season length and bag limit.

Image from MexFish.com

We know that how people respond in surveys is often not a reliable indication of how they will actually make choices. Therefore, we'd like you to respond in this survey as if your decisions are real. Imagine that you actually will have to dig into your pocket and pay the additional trip expenses, and think of the alternative uses of your time. If you choose to take more fishing trips, remember that you will have less time and money to spend on other activities.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

CB Part 1

We are going to present you with two different fishing policy scenarios and would like you to respond with how these policy scenarios would affect the number and timing of your headboat fishing trips in 2015. In thinking of the trips you would take, please treat each scenario completely independently from the other.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

The number of headboat trips you actually took in 2015 is presented below to help you answer this question.

2015 Gulf of Mexico Headboat Trips you Actually Took

	January to May	June	July to end of August	September to end of December
Number of partial day (4-8 hrs) headboat trips in 2015 actually taken	${e://Field/part_trips1}	${e://Field/part_trips2}	${e://Field/part_trips3}	${e://Field/part_trips4}
Number of full day (8-15 hrs) headboat trips in 2015 actually taken	${e://Field/full_trips1}	${e://Field/full_trips2}	${e://Field/full_trips3}	${e://Field/full_trips4}

In recent years recreational anglers could only retain ${e://Field/Fish} during a ${e://Field/Timetext}. The season length and bag limit of a typical ${e://Field/Fish} season in the recent past are presented in Policy A below.

Policy A

Season when ${e://Field/Fish} can be retained	${e://Field/Season}

${e://Field/FishCap} bag limit	2
Price of one partial day (4-8 hrs) headboat trip	$${e://Field/cb_partialcostA}
Price of one full day (8-15 hrs) headboat trip	$${e://Field/cb_fullcostA}

If the Gulf of Mexico ${e://Field/Fish} fishing policies were as described in Policy A, how many headboat trips would you have taken in 2015 in the different seasons? In considering your responses, please assume that any features about the fishing trips that are not mentioned such as sea conditions, the quality and size of the boat, the number of passengers, and bag limits and regulations for other species are the same as your 2015 experience.

2015 Gulf of Mexico Headboat Trips under Policy A

	January to end of May Holidays: New Year's, Spring break, Easter, Memorial day	June	July to end of August Holidays: Independence Day	September to end of December Holidays: Labor Day, Columbus Day, Thanksgiving, Christmas
Number of partial day (4-8 hrs) headboat trips in 2015 under Policy A				
Number of full day (8-15 hrs) headboat trips in 2015 under Policy A				

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Now consider an alternative policy, Policy B. The number of headboat trips you actually took in 2015 is presented below to help you answer this question.

2015 Gulf of Mexico Headboat Trips you Actually Took

	January to May	June	July to end of August	September to end of December
Number of partial day (4-8 hrs) headboat trips in 2015 actually taken	${e://Field/part_trips1}	${e://Field/part_trips2}	${e://Field/part_trips3}	${e://Field/part_trips4}
Number of full day (8-15 hrs) headboat trips in 2015 actually taken	${e://Field/full_trips1}	${e://Field/full_trips2}	${e://Field/full_trips3}	${e://Field/full_trips4}

Please assume that any features about the fishing trips that are not mentioned such as sea conditions, the quality and size of the boat, the number of passengers, and bag limits and regulations for other species are the same as your 2015 experience.

Policy B

Season when ${e://Field/Fish} can be retained	Any time of year
${e://Field/FishCap} bag limit	${e://Field/cb_bagB}
Price of one partial day (4-8 hrs) headboat trip	$${e://Field/cb_partialcostB}
Price of one full day (8-15 hrs) headboat trip	$${e://Field/cb_fullcostB}

If the Gulf of Mexico ${e://Field/Fish} fishing policies were as described in Policy B, how many headboat trips would you have taken in 2015 in the different seasons?

2015 Gulf of Mexico Headboat Trips under Policy B

	January to end of May Holidays: New Year's, Spring break, Easter, Memorial day	June	July to end of August Holidays: Independence Day	September to end of December Holidays: Labor Day, Columbus Day, Thanksgiving, Christmas
Number of partial day (4-8 hrs) headboat trips in 2015 under Policy B				
Number of full day (8-15 hrs) headboat trips in 2015 under Policy B				

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

CB Part 2

Taking into account your responses to the previous questions, overall which type of fishing season would you prefer?

	Policy A	Policy B
Season when ${e://Field/Fish} can be retained	${e://Field/Season}	Any time of year
${e://Field/FishCap} bag limit	2	${e://Field/cb_bagB}
Price of one partial day (4-8 hrs) headboat trip	$${e://Field/cb_partialcostA}	$${e://Field/cb_partialcostB}

Price of one full day (8-15 hrs) headboat trip	$${e://Field/cb_fullcostA}		$${e://Field/cb_fullcostB}	
Number of trips you would take per season in 2015	${e://Field/ptrips1A} partial day ${e://Field/ftrips1A} full day	January to May	${e://Field/ptrips1B} partial day ${e://Field/ftrips1B} full day	January to May
	${e://Field/ptrips2A} partial day ${e://Field/ftrips2A} full day	June	${e://Field/ptrips2B} partial day ${e://Field/ftrips2B} full day	June
	${e://Field/ptrips3A} partial day ${e://Field/ftrips3A} full day	July to August	${e://Field/ptrips3B} partial day ${e://Field/ftrips3B} full day	July to August
	${e://Field/ptrips4A} partial day ${e://Field/ftrips4A} full day	September to December	${e://Field/ptrips4B} partial day ${e://Field/ftrips4B} full day	September to December

I would prefer...

Policy A

Policy B

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

CE full day - bag limits

We would now like you to consider the decision of whether to take a specific full day (8-15 hrs) headboat fishing trip. We will ask you to look at 4 trip decisions that are all a little different in terms of the following features:

Feature	Description
Total expected number of ${e://Field/Fish} caught per trip	Your expected total catch of ${e://Field/Fish}.
Bag limit	The number of ${e://Field/Fish} that you are legally allowed to keep per fishing trip.
Total number of other species caught per trip	Your expected total catch besides ${e://Field/Fish}
Congestion	Crowded: 2.5 feet between fisherman along the side rails on the boat and boat at full capacity. Spacious: 6 feet between fishermen along the side rails on the boat and boat at 1/2 capacity.
Price for full day trip	The per person cost of the full day (8-15 hrs) headboat trip you pay the headboat company (not including tips).

For the following questions, assume that the trip occurs during a time of year in which you would typically go headboat fishing.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each full day (8-15 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce1_target11} ${e://Field/Fish}	${e://Field/ce1_target12} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Bag limit	${e://Field/ce1_bag11}	${e://Field/ce1_bag12}	
Number of other species caught per trip	${e://Field/ce1_other11} fish	${e://Field/ce1_other12} fish	
Congestion	${e://Field/ce1_congest11}	${e://Field/ce1_congest12}	
Price for full day trip	$${e://Field/ce1_fullprice11}	$${e://Field/ce1_fullprice12}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each full day (8-15 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce1_target21} ${e://Field/Fish}	${e://Field/ce1_target22} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Bag limit	${e://Field/ce1_bag21}	${e://Field/ce1_bag22}	
Number of other species caught per trip	${e://Field/ce1_other21} fish	${e://Field/ce1_other22} fish	
Congestion	${e://Field/ce1_congest21}	${e://Field/ce1_congest22}	
Price for full day trip	$${e://Field/ce1_fullprice21}	$${e://Field/ce1_fullprice22}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Use of Funds A

Many recreational fisheries, including ${e://Field/Fish}, are managed through bag limits to help ensure the fishery is not depleted. An alternative management option used in some fisheries is where fishermen pay a fee per fish they retain. For the next two choices, assume that there is an alternative fishery management in place where there are no limits on the number of ${e://Field/Fish} you can retain (i.e. no bag limits), but rather a fee for each ${e://Field/Fish} retained.

The fees would be collected by the headboat operators as customers leave the vessel at port. The headboat operators would keep the money from the fees in return for operating under strict regulatory limits per vessel on the number of ${e://Field/Fish} that their customers can catch and keep for the year.

How acceptable do you find the fishery management option where there are no limits on the number of ${e://Field/Fish} you can retain (i.e. no bag limits), but rather a fee for each ${e://Field/Fish} retained?

	Definitely Acceptable	Somewhat Acceptable	Neither Acceptable nor Unacceptable	Somewhat Unacceptable	Definitely Unacceptable
Management option with a fee for each ${e://Field/Fish} retained	○	○	○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Use of Funds B

Many recreational fisheries, including ${e://Field/Fish}, are managed through bag limits to help ensure the fishery is not depleted. An alternative management option used in some fisheries is where fishermen pay a fee per fish they retain. For the next two choices, assume that there is an alternative fishery management in place where there are no limits on the number of ${e://Field/Fish} you can retain (i.e. no bag limits), but rather a fee for each ${e://Field/Fish} retained.

The fee would be collected by the headboat operators as people leave the vessel at port. The money collected by the headboat operators would be used to fund habitat enhancement projects in the Gulf of Mexico and Gulf of Mexico fishery research.

How acceptable do you find the fishery management option where there are no limits on the number of ${e://Field/Fish} you can retain (i.e. no bag limits), but rather a fee for each ${e://Field/Fish} retained?

	Definitely Acceptable	Somewhat Acceptable	Neither Acceptable nor Unacceptable	Somewhat Unacceptable	Definitely Unacceptable
Management option with a fee for each ${e://Field/Fish} retained	○	○	○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

CE full day - fee per fish

Please examine the table below, compare all the features of each full day (8-15 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip

Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce2_target11} ${e://Field/Fish}	${e://Field/ce2_target12} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Cost per each retained red snapper	$${e://Field/ce2_fee11}	$${e://Field/ce2_fee12}	
Number of other species caught per trip	${e://Field/ce2_other11} fish	${e://Field/ce2_other12} fish	
Congestion	${e://Field/ce2_congest11}	${e://Field/ce2_congest12}	
Price for full day trip	$${e://Field/ce2_fullprice11}	$${e://Field/ce2_fullprice12}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee11} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target11} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee12} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target12} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each full day (8-15 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce2_target21} ${e://Field/Fish}	${e://Field/ce2_target22} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Cost per each retained red snapper	$${e://Field/ce2_fee21}	$${e://Field/ce2_fee22}	
Number of other species caught per trip	${e://Field/ce2_other21} fish	${e://Field/ce2_other22} fish	
Congestion	${e://Field/ce2_congest21}	${e://Field/ce2_congest22}	
Price for full day trip	$${e://Field/ce2_fullprice21}	$${e://Field/ce2_fullprice22}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee21} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target21} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee22} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target22} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

CE partial day - bag limits

We would now like you to consider the decision of whether to take a specific partial day (4-8 hrs) headboat fishing trip. We will ask you to look at 4 trip decisions that are all a little different in terms of the following features:

Feature	Description
Total expected number of ${e://Field/Fish} caught per trip	Your expected total catch of ${e://Field/Fish}.
Bag limit	The number of ${e://Field/Fish} that you are legally allowed to keep per fishing trip.
Total number of other species caught per trip	Your expected total catch besides ${e://Field/Fish}
Congestion	Crowded: 2.5 feet between fisherman along the side rails on the boat and boat at full capacity. Spacious: 6 feet between fishermen along the side rails on the boat and boat at 1/2 capacity.
Price for partial day trip	The per person cost of the partial day (4-8 hrs) headboat trip you pay the headboat company (not including tips).

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each partial day (4-8 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover over the feature to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce1_target11} ${e://Field/Fish}	${e://Field/ce1_target12} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Bag limit	${e://Field/ce1_bag11}	${e://Field/ce1_bag12}	
Number of other species caught per trip	${e://Field/ce1_other11} fish	${e://Field/ce1_other12} fish	
Congestion	${e://Field/ce1_congest11}	${e://Field/ce1_congest12}	
Price for partial day trip	$${e://Field/ce1_partprice11}	$${e://Field/ce1_partprice12}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each partial day (4-8 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover over the feature to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce1_target21} ${e://Field/Fish}	${e://Field/ce1_target22} ${e://Field/Fish}	Do something else, but do not go

			saltwater fishing on a headboat
Bag limit	${e://Field/ce1_bag21}	${e://Field/ce1_bag22}	
Number of other species caught per trip	${e://Field/ce1_other21} fish	${e://Field/ce1_other22} fish	
Congestion	${e://Field/ce1_congest21}	${e://Field/ce1_congest22}	
Price for partial day trip	$${e://Field/ce1_partprice21}	$${e://Field/ce1_partprice22}	

I would choose...

Trip 1 ○ Trip 2 ○ No Trip ○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

CE partial day - fee per fish

Please examine the table below, compare all the features of each partial day (4-8 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip
Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce2_target11} ${e://Field/Fish}	${e://Field/ce2_target12} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Cost per each retained red snapper	$${e://Field/ce2_fee11}	$${e://Field/ce2_fee12}	
Number of other species caught per trip	${e://Field/ce2_other11} fish	${e://Field/ce2_other12} fish	

Congestion	${e://Field/ce2_congest11}	${e://Field/ce2_congest12}	
Price for partial day trip	$${e://Field/ce2_partprice11}	$${e://Field/ce2_partprice12}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee11} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target11} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee12} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target12} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Please examine the table below, compare all the features of each partial day (4-8 hrs) fishing trip, and then answer the question below as if these two headboat trips were the only ones available to you. Please consider the timing of the two trips as identical, occurring at a time of year when you most want to take a fishing trip.

Hover your cursor over the feature name to see its definition.

Features	Trip 1	Trip 2	No trip

Total expected number of ${e://Field/Fish} caught per trip	${e://Field/ce2_target21} ${e://Field/Fish}	${e://Field/ce2_target22} ${e://Field/Fish}	Do something else, but do not go saltwater fishing on a headboat
Cost per each retained red snapper	$${e://Field/ce2_fee21}	$${e://Field/ce2_fee22}	
Number of other species caught per trip	${e://Field/ce2_other21} fish	${e://Field/ce2_other22} fish	
Congestion	${e://Field/ce2_congest21}	${e://Field/ce2_congest22}	
Price for partial day trip	$${e://Field/ce2_partprice21}	$${e://Field/ce2_partprice22}	

I would choose...

Trip 1	Trip 2	No Trip
○	○	○

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee21} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target21} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Under your preferred trip it costs $${e://Field/ce2_fee22} to keep one ${e://Field/Fish}. If you caught ${e://Field/ce2_target22} ${e://Field/Fish}, how many ${e://Field/Fish} would you pay the fee to retain?

Number of ${e://Field/Fish} kept

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds

Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

CB-CE Intro Red Snapper

Managing the Gulf of Mexico Fisheries

Fisheries managers are considering changes to recreational fishing policy for headboats in the Gulf of Mexico that may impact ${e://Field/Fish} season length and bag limit.

Image from Flyfishingpoint.net

We know that how people respond in surveys is often not a reliable indication of how they will actually make choices. Therefore, we'd like you to respond in this survey as if your decisions are real. Imagine that you actually will have to dig into your pocket and pay the additional trip expenses, and think of the alternative uses of your time. If you choose to take more fishing trips, remember that you will have less time and money to spend on other activities.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Time valuation working

We are now going to ask you some questions about how flexible your schedule is and how you choose between working and participating in recreation activities like headboat fishing.

How many hours do you work in a typical week?

Number of hours

Do you work a fixed hourly schedule, such as 9 to 5 Monday through Friday or the same shifts every week, or are you free to choose when and how long you work?

- Fixed hourly schedule
- Free to choose when and how long to work

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

If you had the opportunity to:

work more days, receive more income, but have less time for leisure activities

OR

work fewer days, receive less income, but have more time for leisure activities, which option would you take?

Please choose one of these three options.

Work More	Work Less	No Change
• Earn more income • Have less leisure time	• Earn less income • Have more leisure time	• Same hours worked and income as current situation

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

WTP question

Some companies allow employees to take an additional day off work during the summer by giving up some money on their next paycheck.

Suppose you could give up some earnings off one of your paychecks to receive an additional day off (8 hours) during the three summer months (June, July and August).

How likely is it you would you take one additional day off work (8 hours) if you had to give up…? Please select a response for each payroll deduction amount.

	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
$50					
$100					
$200					
$400					
$700					

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

In the previous question, why did you respond 'Definitely No' to all the payroll deduction amounts for additional time off work?

- I do not think it is worth the money
- I do not believe that my company would offer such an option
- I do not think it is appropriate to have to give up money for additional time off

WTA focus group

Many research companies pay people to participate in a focus group.

Suppose you’ve been given the opportunity to be paid a certain amount of money to participate in a full day (8 hours) focus group near your home during one of your days off during the three summer months (June, July and August).

How likely is it you would participate in the focus group if the payment amount is…? Please select a response for each payment amount.

	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
$50					
$100					
$200					
$400					
$700					

Definitely Yes (100% chance) Probably Yes (75% chance) Not Sure (50% chance) Probably No (25% chance) Definitely No (0% chance)

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

WTA contract

Many companies look to hire people on a short term contract for one or two days.

Suppose you’ve been given the opportunity to be paid a certain amount of money to work for one day (8 hours) for a company sorting paper files alphabetically. The work would be near your home during one of your days off during the three summer months (June, July and August).

How likely is it you would participate in the short term contract if the payment amount is…? Please select a response for each payment amount.

	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
$50					
$100					
$200					
$400					
$700					

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Leisure tradeoffs

We are now going to ask you to choose between two options that change the number of leisure days between different seasons of the year.

Suppose you were given two options relative to your current situation in terms of the timing of your leisure days. Which of the options below do you most and least prefer?

Time period	Option A	Option B	Option C
January to May	${e://Field/winter1a}	${e://Field/winter1b}	No change from your current situation
June to August	1 less leisure day	1 less leisure day	
September to December	${e://Field/fall1a}	${e://Field/fall1b}	

	Option A	Option B	Option C
I most prefer...			
I least prefer...			

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds

Last Click: 0 seconds

#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Now suppose you were given two options relative to your current situation in terms of the timing of your leisure days. Which of the options below do you most and least prefer?

Time period	Option A	Option B	Option C

January to May	${e://Field/winter2a}	${e://Field/winter2b}	No change from your current situation
June to August	1 less leisure day	1 less leisure day	
September to December	${e://Field/fall2a}	${e://Field/fall2b}	

	Option A	Option B	Option C
I most prefer...	○	○	○
I least prefer...	○	○	○

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Time valuation not-working

We are now going to ask you some questions about how you choose between earning additional income and participating in recreation activities like headboat fishing.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Demographics

Now we just have a few more questions to ask you that will help us understand your responses compared to other members of the public.

In what year were you born (enter 4-digit birth year; for example, 1973)

4 - digit birth year

Are you male or female?

○ Male

○ Female

How many individuals live in your household?

Enter Number of Household Members

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds

Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

How many children under the age of 18 live in your household?

Enter number of children

What is the highest level of school you have completed or the highest degree you have received?

- Less than high school degree
- High school degree or equivalent (e.g., GED)
- Some college but no degree
- Associate degree
- Bachelor degree
- Graduate degree

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Are you a member or do you contribute to any of the following classes of organizations? Please select all that apply.

- Environmental advocacy organizations (Sierra Club, Defenders of Wildlife, etc.)
- Recreational angler associations (Coastal Conservation Association, International Game Fish Association, etc.)
- Hunting organizations (Ducks Unlimited, Whitetails Unlimited, etc.)

What was your household's total income before taxes in 2015?

Less than $24,999	$75,000 - $99,999	$150,000 - $199,999
$25,000 - $49,999	$100,000 - $124,999	$200,000 - $249,999
$50,000 - $74,999	$125,000 - $149,999	Greater than $250,000

Think about how much you and your household spent on vacation and recreation activities outside of the house in all of 2015. Please think about vacations and trips you took, travel, and expenses on recreation activities. Roughly, how much would that amount to?

Dollars spent on vacation and recreation activities outside of the home in 2015

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds

#QuestionText, TimingClickCount#: 0 clicks

Thank you for your responses to the survey!

Please add any additional comments you may have about this survey in the space provided.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds
Last Click: 0 seconds
#QuestionText, TimingPageSubmit#: 0 seconds
#QuestionText, TimingClickCount#: 0 clicks

Supplementary Material C: Public Good WTA Experiment Scripts

[Public Good Real Treatment]

Instructions

In this experiment, you have the opportunity as a group to collectively vote on what to do with the last 30 minutes of time in this session. If a majority of people in this room vote yes then everyone will be paid an additional $3 and will have to give up the 30 minutes and fill out a questionnaire. If a majority of people in this room vote no then everyone can leave immediately after the vote and do not have to fill out the questionnaire, but will not receive the additional $3 each.

Below please find the proposition and referendum rules.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Referendum Rules:

- If more than 50% of you vote YES on this proposition, all of you will receive an additional $3. In exchange, each participant must give up 30 minutes of their time and fill out a questionnaire.

- If 50% or fewer of you vote YES on this proposition, no one will receive an additional $3 and no one will have to give up 30 minutes of their time and fill out a questionnaire.

Are there any questions? Please turn over to your decision sheet.

[Public Good Real Treatment]

<u>Decision Sheet</u>

The following situation involves real monetary impacts, and the outcome, based on the decisions of you and the other respondents, will be binding.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Referendum Rules:

- If more than 50% of you vote YES on this proposition, all of you will receive an additional $3. In exchange, each participant must give up 30 minutes of their time and fill out a questionnaire.

- If 50% or fewer of you vote YES on this proposition, no one will receive an additional $3 and no one will have to give up 30 minutes of their time and fill out a questionnaire.

Do you vote YES or NO in the referendum? That is, do you vote for or against having all participants in this room receive an additional $3 to give up 30 minutes of time to fill out a questionnaire?

- ❑ Vote **YES,** I am willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

- ❑ Vote **NO,** I am not willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

[Public Good Consequential Treatment]

Instructions

In this experiment, you have the opportunity as a group to collectively vote on what to do with the last 30 minutes of time in this session. If a majority of people in this room vote yes to the proposition, then there is a chance everyone will be paid an additional $3 and will have to give up the 30 minutes and fill out a questionnaire. If a majority of people in this room vote no then everyone can leave immediately after the vote and do not have to fill out the questionnaire, but will not receive the additional $3 each.

Below please find the proposition and referendum rules.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Two-Step Referendum Rules:

Step 1:

- If more than 50% of you vote YES on this proposition, then the referendum has **passed**. If the referendum passes, then in Step 2 a coin will be flipped to determine if the referendum is binding.

- If 50% or fewer of you vote YES on this proposition, then the referendum does not pass and no one will receive an additional $3. Hence, no one will have to give up 30 minutes of their time and fill out a questionnaire.

Step 2:

- Contingent on the referendum passing (more than 50% of you vote YES), a coin will be flipped [show coin]. If the coin lands heads, the referendum will be binding and all of you will receive an additional $3 and will have to give up the 30 minutes and fill out a questionnaire. If the coin lands tails, the referendum is not binding. In this case, no one receives an additional $3 and no one has to fill out the questionnaire.

Are there any questions? Please turn over to your decision sheet.

[Public Good Consequential Treatment]

<u>Decision Sheet</u>

The following situation involves real monetary impacts, and the outcome, based on the decisions of you and the other respondents, will be binding.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Two-Step Referendum Rules:

Step 1:

- If more than 50% of you vote YES on this proposition, then the referendum has **passed**. If the referendum passes, then in Step 2 a coin will be flipped to determine if the referendum is binding.

- If 50% or fewer of you vote YES on this proposition, then the referendum does not pass and no one will receive an additional $3. Hence, no one will have to give up 30 minutes of their time and fill out a questionnaire.

Step 2:

- Contingent on the referendum passing (more than 50% of you vote YES), a coin will be flipped [show coin]. If the coin lands heads the referendum will be binding and all of you will receive an additional $3 and will have to give up the 30 minutes and fill out a questionnaire. If the coin lands tails, the referendum is not binding. In this case, no one receives an additional $3 and no one has to fill out the questionnaire.

Do you vote YES or NO in the referendum? That is, do you vote for or against having all participants in this room receive an additional $3 to give up 30 minutes of time to fill out a questionnaire?

- ❑ Vote **YES,** I am willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

- ❑ Vote **NO,** I am not willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

[Public Good Hypothetical Treatment]

Instructions

In this experiment, suppose you have the opportunity as a group to collectively vote on what to do with the last 30 minutes of time in this session. If a majority of people in this room vote yes to the proposition, then everyone would be paid an additional $3 and would have to give up the 30 minutes and fill out a questionnaire. If a majority of people in this room vote no then everyone can leave immediately after the vote and do not have to fill out the questionnaire, but will not receive the additional $3 each.

Regardless of the vote outcome, no one will receive an additional $3 or have to stay the extra 30 minutes to fill out the questionnaire. We ask that you vote as if you were actually making a real money decision on whether to accept the additional $3 and give up the 30 minutes of your time.

Below please find the proposition and referendum rules.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Referendum Rules:

- If more than 50% of you vote YES on this proposition, all of you will receive an additional $3. In exchange, each participant must give up 30 minutes of their time and fill out a questionnaire.

- If 50% or fewer of you vote YES on this proposition, no one will receive an additional $3 and no one will have to give up 30 minutes of their time and fill out a questionnaire.

Are there any questions? Please turn over to your decision sheet.

[Public Good Hypothetical Treatment]

Decision Sheet

The following situation does not involve real monetary impacts or outcomes. We ask that you vote as if you were actually making a real money decision.

Proposition: Everyone in the room will be paid an additional $3. In exchange, each participant will have to give up 30 minutes and fill out a questionnaire.

Referendum Rules:

- If more than 50% of you vote YES on this proposition, all of you will receive an additional $3. In exchange, each participant must give up 30 minutes of their time and fill out a questionnaire.

- If 50% or fewer of you vote YES on this proposition, no one will receive an additional $3 and no one will have to give up 30 minutes of their time and fill out a questionnaire.

Would you vote YES or NO in the referendum? That is, do you vote for or against having all participants in this room receive an additional $3 to give up 30 minutes of time to fill out a questionnaire?

- ☐ Vote **YES,** I am willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

- ☐ Vote **NO,** I am not willing to accept an additional $3 to give up 30 minutes of my time to fill out a questionnaire.

Supplementary Material D: Private Good WTA Experiment Scripts

Instructions *[The main text presents the consequential treatment. Differences between treatments are provided in the parentheses.]*

Thank you for participating in today's experiment.

Your participation is valuable and helps us learn about economic behavior and decision making. During the experiment, you will have opportunities to earn additional money. I will use a script throughout the experiment to ensure that I communicate information and instructions consistently. However, feel free to ask any questions that you have. Just raise your hand at any time. It is important that when we begin no one talks to anyone but a moderator. Are we ready?

Experiment Description

In this first experiment, you will be asked questions about whether you would be willing to be paid money to give up one hour of your time at a time and date of your choosing. During this hour you will be organizing books in the Resource Economics and Environmental Sociology department's library. This hour of time can be completed on any weekday between 9 am and 5 pm and must be completed within two weeks of today's date.

[Hypothetical Treatment: You will not actually have to give up the time nor will you receive the money. But we ask that you make choices as if you were actually making a real money decision on whether to accept the payments and give up an hour of your time.]

This experiment is divided into two stages: an initial **Survey Stage** and then an **Offer Stage**.

I will first describe the **Offer Stage** because you will be making a real monetary offer *[Hypothetical Treatment: an offer]* of the minimum dollar amount you would be willing to accept to give up an hour of your time as described earlier. Your offer will be compared to a randomly drawn payment level contained in one of the envelopes held by the moderator. There are two possible outcomes:

1. **Your offer is accepted**. If the amount you indicate is less than or equal to the randomly drawn payment level, then you will have to give up the hour of your time and you will receive the randomly drawn payment level in dollars.

2. **Your offer is not accepted.** If the amount you indicate is more than the randomly drawn payment level, then you will not have to give up the hour of time and you do not receive any money.

We are running several of these experiments with different groups of people, but we have a limited budget to pay for help in the library. Thus, we will not be able to provide the **Offer Stage** to every group. In addition, we may provide different payment level envelopes to different groups *[Price Treatment: Thus, we may* provide *different payment level envelopes to different groups.] [Provision Treatment: Thus, we will not be able to provide the **Offer Stage** to every group.]*.

Before the **Offer Stage**, in the **Survey Stage** you will be asked to answer how likely you would be to accept different payment levels to give up one hour of your time as described earlier.

Survey Stage

We are now going to begin the **Survey Stage** of the experiment.

Responses in the **Survey Stage** will help the researchers decide whether or not to present the **Offer Stage** to this group as well as determine which payment level envelopes to provide in the **Offer Stage**, depending on our budget.

*[Price Treatment: Responses in the **Survey Stage** will help the researchers determine which payment level envelopes to* provide *in the **Offer Stage**, depending on our budget.]*

[Provision Treatment: *Responses in the **Survey Stage** will help the researchers decide whether or not to* present *the **Offer Stage** to this group, depending on our budget. The payment levels potentially being provided in the **Offer Stage** have already been determined. (show envelopes)]*

However, the answers in this stage will not be used to determine if you will accept the payments for your time or not.

Any questions?

Please turn to the Survey Stage sheet on the next page.

Survey Stage *[The next pages are handed out to the participants separately]*

In responding to the following question, think about how your answers will help the researchers decide whether to present the **Offer Stage** to this group or not as well as determine which payment level envelopes to provide in the **Offer Stage**, depending on our budget.

[Price Treatment: In responding to the following question, think about how your answers will help the researchers determine which payment level envelopes to provide *in the **Offer Stage**, depending on our budget.]*

*[Provision Treatment: In responding to the following question, think about how your answers will help the researchers decide whether or not to present the **Offer Stage** to this group, depending on our budget. The payment levels potentially being provided in the **Offer Stage** have already been determined. (show envelopes)]*

The choices you make in this stage will not be used to determine if you will accept the payments or not. Please respond to the following questions thoughtfully and honestly.

Suppose you have been given the opportunity to be paid money to give up one hour of your time to work in the department's library helping sort and organize library books. The work can be completed any weekday between 9 am and 5 pm and the work must be completed within two weeks of today's date (March 26, 2017). Please think about a suitable time when you would complete the work before you respond to the question. You would be paid once you've completed the hour of time.

How likely would you be to accept the offer to give up one hour of your time if the payment amount is....? Please select a response for each payment amount (one response per row).

Payment for 1 Hour of Time	Definitely Yes (100% chance)	Probably Yes (75% chance)	Not Sure (50% chance)	Probably No (25% chance)	Definitely No (0% chance)
$5	□	□	□	□	□
$10	□	□	□	□	□
$15	□	□	□	□	□
$20	□	□	□	□	□
$25	□	□	□	□	□
$30	□	□	□	□	□
$35	□	□	□	□	□
$40	□	□	□	□	□

Once you have completed the Survey Stage sheet handed to you, please raise your hand and a moderator will collect your responses and give you an additional set of questions.

Which of the following motives were important when deciding to accept or not accept the offer at each payment level? *Select all that apply.*

- ☐ A. Ensure the offer is worth it for myself given the payment levels presented
- ☐ B. Ensure the Offer Stage is provided to my group
- ☐ C. Ensure the payment levels in the Offer Stage are favourable for myself

If you selected more than one motive, which one is the most important? *Please write the letter associated with the motive below.*

____

On a scale from 1 to 5, where 1 is not at all and 5 is definitely taken into account, to what extent do you think your choices will be taken into account for determining **which payment levels are provided** in the **Offer Stage**? *Please circle the number corresponding to your response.*

	Not taken into account		Somewhat taken into account		Definitely taken into account
Which payment levels are provided in the Offer Stage	1	2	3	4	5

On a scale from 1 to 5, where 1 is not at all and 5 is definitely taken into account, to what extent do you think your choices will be taken into account for determining the **chances of being provided** the **Offer Stage**? *Please circle the number corresponding to your response.*

	Not taken into account		Somewhat taken into account		Definitely taken into account
The chances of being provided the Offer Stage	1	2	3	4	5

Once you have completed these questions please wait while the researchers examine the Survey Stage results and decide whether to present the **Offer Stage** to this group or not as well as determine which payment level envelopes to provide in the **Offer Stage**,

[Price Treatment: Once you have completed these questions please wait while the researchers examine the Survey Stage results and decide which payment level envelopes to provide *in the* ***Offer Stage****,*

[Provision Treatment: Once you have completed these questions please wait while the researchers examine the Survey Stage results and decide whether to present the ***Offer Stage*** *to this group or not.]*

Offer Stage

We are now going to begin the **Offer Stage** of the experiment.

Responses in this Offer Stage are real and have monetary and time consequences.

[Hypothetical Treatment: You will not actually have to give up the time nor will you receive the money. But we ask that you make choices as if you were actually making a real money decision on whether to accept the payments and give up an hour of your time.]

The moderator is willing to buy an hour of your time from you. Below, please make an offer and indicate the minimum amount you would be willing to accept to sell this hour of time to the moderator. During the hour you will be organizing books in the department's library and this work can be completed any weekday between 9 am and 5 pm within two weeks of today's date.

There are 10 payment levels in different envelopes on the table. One of the envelopes will be selected at random. There are two possible outcomes:

1. **Your offer is accepted**. If the amount you indicate below is less than or equal to the randomly drawn payment level, then you will have to give up the hour of your time and you will receive the randomly drawn payment level in dollars.

2. **Your offer is not accepted.** If the amount you indicate below is more than the randomly drawn payment level, then you will not have to give up the hour of time and you do not receive any money.

Notice the following two things:

1. Your decision will have no effect on the payment level actually used because the payment level will be selected at random.
2. It is in your interest to indicate your true preferences in terms of your willingness to accept for the sale of an hour of your time.

After all decision forms have been completed, the moderator will select an envelope at random with the payment level that will determine if your offer is accepted or not.

If your offer is accepted, you will have 48 hours to send an email to the email address on the piece of paper beside you with the times you are available. You will be paid once you've completed the hour of time.

The minimum amount that I would be willing to accept to sell an hour of my time is $________

Now we just have a few more questions to ask you that will help us understand your responses.

1. In what year were you born (enter 4-digit birth year; for example, 1973)

_____

2. What is your gender?

___ Male
___ Female
___ Prefer not to say

3. If you are a student, what is your academic major?

____

4. Do you currently volunteer your time with any organization?

___ Yes
___ No

5. What is your current working situation?

___ Part-time
___ Full-time
___ I do not currently work

6. If you do work, what is your approximate hourly wage?

$___ per hour

7. Besides rent or costs for housing, about how much money do you spend in a typical month?

$_____

8. In this experiment, you were offered the opportunity to earn additional money and give up your time. Given your schedule over the next two weeks, are you able to give up one hour of your time on Monday to Friday between 9 am and 5 pm.

___ Yes
___ No

Supplementary Material E: Water Reliability Survey

Water Management in Cities and Towns

(Note: Images for Brock University and University of Alberta logos appeared here for survey but are omitted for Thesis deposit)

A research project to support policy making and decision making. Sponsored by the Water Economics Policy and Governance Research Network. Conducted by researchers from the University of Alberta and Brock University.

Study Overview

Water Management in Cities and Towns

Principal Investigators:
Vic Adamowicz, Professor at University of Alberta, Department of Resource Economics and Environmental Sociology; Phone: 780-492-4603

Pat Lloyd-Smith, Graduate Student at University of Alberta, Department of Resource Economics and Environmental Sociology; and

Diane Dupont, Professor at Brock University; Department of Economics

James Price, Post-Doctoral Fellow, Brock University, Department of Economics

Diane.dupont@brocku.ca; vic.adamowicz@ualberta.ca; lloydsmi@ualberta.ca; jprice@brocku.ca

You are invited to participate in a study on water management that involves researchers in Alberta and Ontario.

Who is funding this? This study is being funded by the Water Economics, Policy and Governance Network, a network of Canadian researchers who have joined together to look at water issues, and the Canadian Water Network, a multidisciplinary water research and knowledge mobilization network.

Partners in this project include Alberta Environment and Sustainable Resource Development and the Canadian Forest Service.

[PROGRAMMER NOTE: Split sample on inclusion of partner funding note above]

What is the Purpose? The goal of this research is to determine public preferences for improved water management and quality and to avoid adverse outcomes associated with drinking water supply.

What Methods are Being Used? We are asking you to take part in a survey being held across Alberta. This information could be used to structure more efficient water management and pricing schemes for municipal water utilities and to aid these utilities in their infrastructure investment decisions. The survey should take about 25 minutes of your time.

What are the Benefits to You? Survey participants will assist the researchers in obtaining estimates of the public's perceptions of water supplies and quality and the importance of clean and reliable water for Albertans. There are no known or anticipated risks associated with participation in this study.

Confidentiality: All information you provide is considered confidential and grouped with responses from other participants. Names will not be associated with survey responses. Access to the data will be restricted to investigators.

Withdrawal: Participation in this study is voluntary. If you wish, you may decline to answer any questions or participate in any component of the study. Further, you may decide to withdraw from this study at any time and may do so without any penalty or loss of benefits to which you are entitled. Once the survey has been completed you cannot withdraw the information you provided.

Publication of Results: Grouped results of this study may be published in professional journals and presented at conferences as well as in the graduate student's thesis. Feedback about this study will be available December 2015 from the principal investigators using the contact information provided above.

Contact Information and Ethics Clearance: If you have any questions about this study or require further information, please contact the Principal Investigators using the contact information provided below. This study has been reviewed and received ethics clearance through the Research Ethics Board at the University of Alberta (file #Pro00051054) and through the Research Ethics Board at Brock University (File #14-040). For questions regarding participant rights and ethical conduct of research, contact the University of Alberta's Research Ethics Office at (780) 492-2615. Thank you for your assistance in this research project.

Contact for Further Information: Vic Adamowicz, University of Alberta, Department of Resource Economics and Environmental Sociology; Phone: 780-492-4603; Email: Vic.Adamowicz@ualberta.ca

1. Approximately how many people live in your current community (city, town or village)?

   ___Fewer than 15,000
   ___Between 15,000 and 100,000
   ___More than 100,000

2. How long have you lived in your current community (city, town or village)?

   ______ YEARS

3. Before moving to your current community, where were you living before?

   ______ I have always lived in my current community
   ______ A similar sized community in Alberta
   ______ A smaller sized community in Alberta
   ______ A larger sized community in Alberta
   ______ A community outside of Alberta

4. How long do you plan to live in your current community? Please select one timeframe.

   ______ 0- 5 years
   ______ 6-10 years
   ______ More than 10 years

We would like to know your views on various options for investing public funds. What follows is a list of government programs that are partially paid for by your taxes.

5. In your opinion, how important is it for your municipal government to invest in each of the following? Please use a scale of 1 to 5 where **1** means **not at all important** and **5** means **very important**.

Please select one response for each item

[ACROSS TOP OF GRID]
1 – Not at all important
2
3
4
5 – Very important
[DOWN SIDE OF GRID] [RANDOMIZE ORDER]

Policing services
Food/restaurant safety services
Poverty and social assistance programs
Schools and education
Environmental protection
Clean, reliable water supply
Transportation infrastructure (e.g. roads, bridges)

6. Using a scale of 1 to 5 where **1** means **strongly disagree** and **5** means **strongly agree**, please indicate your agreement or disagreement with the following statements regarding environment and development goals.

Please select one response for each item
[ACROSS TOP OF GRID]
1 – Strongly disagree
2
3
4
5 – Strongly agree

[DOWN SIDE OF GRID] [RANDOMIZE ORDER]
Environmental improvement programs that would be harmful to business should not be carried out
Environmental improvements are fine as long as taxes do not increase
Experts should solve environmental issues and the public should only be educated and informed of the decisions
New technology will solve most environmental problems
In the future, humans will be able to understand and control most natural processes.
Human progress is limited only by technology and not by the environment

We now want to ask you a few questions about water in your home and community.

7. Are you on a city/municipal water system?

_____ YES
_____ NO
_____ Don't Know

If YES, do you pay a water bill?

_____ YES
_____ NO

There are three sources of <u>drinking water used in the home</u> that we want you to think about:

(i) Tap water (either from a well or a municipal source)
(ii) In-home Treated Tap Water (In-home filtration using a tap attachment, container style filtration system, refrigerator attachment or boiling)
(iii) Purchased bottled water (water bottles of any size, purchased from a grocery store or a home delivery service, such as Culligan, Alberta Fresh Springs, Water Pure & Simple, etc.)

8. For the three water sources, please indicate the percentage of water you personally consume at home that comes from each source in any given month. If your answer is zero in any category you must enter 0%.

Water Type	% Consumed
Tap water	
In-home Treated/Filtered Tap Water	
Purchased water (bottled or from home delivery)	
Total (100%)	100%

Water Quality
Now we would like to collect some information from you about the quality of your regular water supply.

9. Which, if any, of the following have you experienced with the tap water in your home over the past year? Please select all that apply.

_____ Rusty colour
_____ Sediment (particles at the bottom of a glass)
_____ Unpleasant smell (e.g., musty, chlorine)
_____ Unpleasant taste (e.g., musty, chlorine)
_____ Hard water / mineral deposits
_____ Pollutants or other contamination
_____ Low water flow/insufficient water pressure
_____ Other ______________________________
_____ None of the above

10. Looking forward five years, do you expect the quality of your tap water at home to be...? Please select one response only.

_____ Worse than today
_____ Same as today

_____ Better than today
_____ Don't know

11. Which of the following statements best reflects your personal opinion about health concerns you might have with the tap water in your home? Please select one response only.

_____ Drinking tap water does not pose a problem for my health or my family's health
_____ Drinking tap water poses a minor problem for my health or my family's health
_____ Drinking tap water poses a moderate problem for my health or my family's health
_____ Drinking tap water poses a serious problem for my health or my family's health

12. Comparing health effects from drinking bottled water (purchased water) to health effects from drinking your home's tap water, do you think that bottled water is…? Please select one.

_____ Much more safe than tap water
_____ A little safer than tap water
_____ About as safe as tap water
_____ A little less safe than tap water
_____ Much less safe than tap water
_____ Don't know/Not sure

13. To the best of your knowledge, have you or anyone in your household ever become sick from drinking any of the following types of water in your home? Select one from each row.

	Yes	No	Don't Know
Tap water			
In-home treated tap water (filtered water)			
Purchased bottled water			

14. For each of the following items that might be present in a household's tap water, please indicate if you have heard about it as a concern with drinking tap water and if any of these items has been a special concern in your community. Please select all that apply for each column.

	Heard About it as a Drinking Water Concern	Drinking Water Concern in My Community	Have not heard about is as a Drinking Water Concern

Microbe – E. coli			
Microbe – Cryptosporidium			
Microbe – Giardia (Beaver Fever)			
Chemical – Fluoride			
Chemical – Trihalomethanes			
Metals – Iron, Lead, Mercury			
Chemical – Pesticides			
Chemical - Pharmaceuticals			

15. Considering each of these contaminants, how much of a health concern do you personally believe each poses in your home's tap water? Please select one for each row.

	No Health Concern	Minor Health Concern	Moderate Health Concern	Serious Health Concern	Don't Know/Uncertain
Microbe – E. coli					
Microbe – Cryptosporidium					
Microbe – Giardia (Beaver Fever)					
Chemical – Fluoride					
Chemical – Trihalomethanes					
Metals – Iron, Lead, Mercury					
Chemical – Pesticides					
Chemical - Pharmaceuticals					

Water Reliability

The remainder of the survey will deal with water reliability issues. Water reliability refers to good quality water being available at any time of day without interruptions**.**

16. Have you experienced any loss of service to the tap water **in your home** in the past year? This can be either a planned or unplanned interruption in water availability or service. Please select all that apply.

_____ Tap water was unavailable (cut off) for some period of time
_____ Boil water advisory was issued
_____ We were unable to obtain tap water for other reasons (e.g. plumbing work in the neighborhood or home, etc.)
_____ We didn't drink the tap water because of smells, colour or some other reason – even though there wasn't an official advisory
_____ There was a water use restriction like a lawn watering restriction or some other public restriction or advisory asking for reduced water use.
_____ We have not experienced any loss of service to our tap water in the past year

PROGRAMMER NOTE: Please code the final response option in Q15 so that it may not be selected in conjunction with any other option.

17. How much of an inconvenience have water outages like the ones described in the previous question been for you? (please select one category)

_____ No inconvenience
_____ Minor inconvenience
_____ Moderate inconvenience
_____ Significant inconvenience

18. Do you keep any "back up" sources of water on hand, specifically so that you will have potable water in the event of a reliability problem with your tap water supply (for example, when there is a boil water advisory or a water outage)?

YES _____
NO _____

Some people keep "back up" sources of water on hand, specifically so that they will have potable water in the event of a reliability problem with their tap water supply (for example, when there is a boil water advisory or a water outage).

19. Please indicate which sources of "back up" water you keep on hand (check all that apply).

If you do not keep "back up" sources of water on hand, please select the final option from the list below.

_____ Bottles of water (e.g. a case of small bottles kept specifically for outages)
_____ Water containers (e.g. a large water container kept specifically for outages)
_____ Equipment for boiling large quantities of water
_____ An in-home water treatment system
_____ Individual water purification system (i.e. Katadyn, LifeStraw)
_____ Rain barrels or other outdoor storage systems
_____ Other water supply alternatives (please specify):___________________
_____ Do not keep "back up" sources of water on hand

20. Approximately how much do spend on these “back up” sources of water (i.e. specifically for water outages) in a year?

$________

21. Looking back over the last 10 years, how many times have the following types of water outage (loss of service) events occurred? For example, if you think you have consistently had about one boil water advisory every 2 years, then this would be 5 events in total over 10 years.

Water Outage Event	Number of events over last 10 years
Expected Water Outage	
Planned water outage (i.e. notice given in advance that tap water will be unavailable for a certain amount of time in the future)	____ events
Unexpected Water Outages	
Short-term unexpected water outage **lasting a few hours but less than a day**	____ events
Longer-term unexpected water outage **lasting at least 1-2 days**	____ events
Boil water advisory	____ events
No official advisory, but we didn’t drink tap water because of smells, colour, or some other reason	____ events

22. Looking forward 10 years and based on your experience and understanding of water management in your community, do you think the number of water outages events will

_____ Increase
_____ Stay the same
_____ Decrease
_____ Don't know/Not sure

23. We are now interested in understanding your expectations in terms of annual percent chances of three specific water outage events occurring over the next 10 years. The following table illustrates the relationship between the number of expected events over the next 10 years and the annual percent chance. Note that if you expect more than ten events of a specific water outage type over the next 10 years, the annual percent chance is still expressed as 100%.

Approximate number of events over next 10 years	Annual percent chance of water outage event over next 10 years
0 in 10 years	0%
1 in 10 years	10%
2 in 10 years	20%
…	…
9 in 10 years	90%
10+ in 10 years	100%

24. To ensure that we have communicated the idea of percent of water outages we would like you to answer the following question.

Suppose you are given the choice of living in one of two communities that are identical except for their annual percent chance of a water outage. Community **A** faces 2 water outages in 10 years, whereas community **B** has a 30% annual chance of water outage over the next 10 years. Which community would you choose to live in? *Please select one response only.*

____Community A: 2 water outages in 10 years
____Community B: 30% annual chance of water outage over the next 10 years.

PROGRAMMER NOTE: If respondent chooses community B, add the following statement

"You answered community B, but that community will have 3 water outages in a ten year period, which is more than community A. The 30% annual chance of water outage means that there will be about 3 outages in 10 years."

25. Looking forward 10 years and based on your experience and understanding of water management in your community, what would be your best guess of the annual percent chance that you (your household) will experience the following water outage events. *Please write your response between 0 and 100 in the following table*

Water Outage Event	Annual Percent Chance of Water Outage Event over next 10 years (0-100%)
24a: A short-term unexpected water outage **lasting a couple of hours**	____%
24b: A longer-term unexpected water outage **lasting at least 1-2 days**	____%
24c: A boil water advisory	____%

26. How confident are you of your responses in the previous question? For each level, please select one response only using a scale from 1 to 5, where 1 is not confident and 5 is confident.

Confidence of expectations	Not Confident		Somewhat Confident		Confident
A short-term unexpected water outage **lasting a few hours but less than a day**	1	2	3	4	5
A longer-term unexpected water outage **lasting at least 1-2 days**	1	2	3	4	5
A boil water advisory	1	2	3	4	5

27. Suppose that you received a letter telling you to expect two water supply interruptions to occur without warning over the next 12 months. You could expect each of these interruptions to be repaired within 1 to 2 days. What action would you take?

Please choose one of the options below.

_____ Take no action to prepare for an interruption (no cost to you)

______ Spend about $5 buying bottled water to keep in the house

______ Spend about $35 buying a 25-litre water container to keep in the house

______ Spend about $70 buying two 25-litre water containers to keep in the house

______ Spend about $105 buying three 25-litre water containers to keep in the house.

Internet Service Outages

The following two questions consider at-home *internet service outages and/or interruptions.* We are interested in comparing your experiences with internet outages to your experiences with water outages.

28. Have you experienced any loss of *internet* services in the past 5 years? This can be either a planned or unplanned interruption in internet service. Please select all that apply.

    ______ The internet was out for some period of time because of an unexpected event (i.e. technical issues, storm)
    ______ The internet was not available because of a planned interruption and the internet company provided advance notice (i.e. system upgrade in local area)
    ______ I have not experienced a loss of internet service in the past 5 years

29. Looking forward 5 years and based on your experience of *internet services* in your community, what would be your best guess of the annual percent chance that you (your household) will experience at least 1 of the following events? *Please write your response in the following table.*

    Remember to treat each internet outage event independently (i.e., please enter a response between 0 and 100 for each row in the following table).

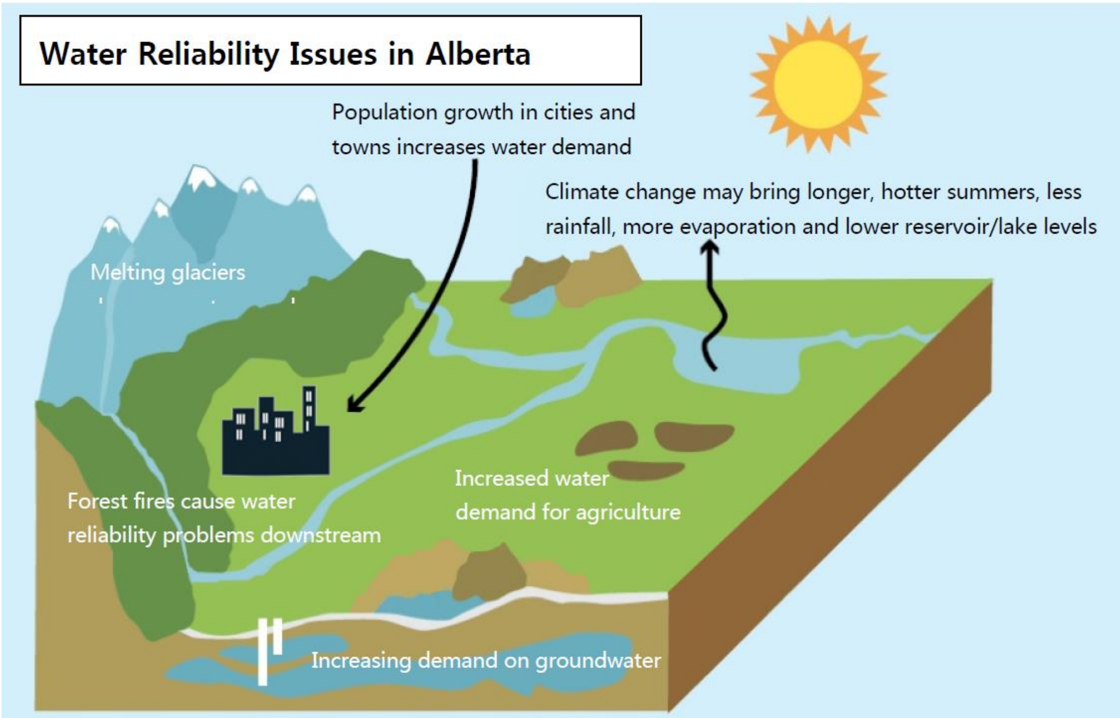
Internet Outage Event	Annual Percent Chance of Internet Outage Event over next 5 years (0-100%)
A short-term unexpected internet outage **lasting a few hours but less than a day**	____%
A longer-term unexpected internet outage **lasting at least 1-2 days**	____%
A series of unexpected outages that come and go sporadically and last minutes rather than hours (i.e., a patchy connection)	____%

THE FUTURE OF DRINKING WATER IN ALBERTA

We would like to turn your attention back to drinking water reliability in Alberta.

Over the last few years many parts of Alberta have experienced water shortages, unplanned outages or boil water advisories. Scientists are concerned that summer droughts will become more frequent and severe in Alberta leading to an increase in the frequency / severity of water shortages. Click here for a graphic illustrating some of the potential water supply and demand issues in the province.

[GRAPHIC FOR HOVERLINK]



Source: Adapted from Prairie Adaptation Research Collaborative (http://www.parc.ca/research_projects-ssrb.htm)

Another water reliability concern relates to forest fires. The vast majority of Alberta's drinking water originates from the forested slopes of the Rocky Mountains. A recent scientific study in southwestern Alberta has documented the effects of forest fires on water quality and the potential for negative downstream impacts on drinking water treatment systems (Emelko et al., 2011). These negative impacts might lead to water outages for communities.

[INCLUDE IN HOVER LINK for Emelko et al., 2010]
Emelko, MB, U. Silins, KD. Bladon, M. Stone. 2011. "Implications of land disturbance on drinking water treatability in a changing climate: Demonstrating the need for "source water supply and protection" strategies". *Water Research* 45:461-472
[END OF HOVER LINK]

30. Are you concerned that summer droughts will become more frequent and severe in Alberta?

_____ Yes
_____ No
_____ Don't know

31. Are you concerned that forest fires will become more frequent and severe in Alberta?

_____ Yes
_____ No
_____ Don't know

Other communities in North America, such as Denver, Colorado, have experienced water reliability problems from droughts and forest fires and have identified various ways to reduce the impact of these events on water reliability. Denver recently implemented additional water fees on households to modify forest vegetation and improve water reliability (click here for more info). These preventative measures might include increasing the capacity of the water treatment plant or modifying the equipment in the plant. Other actions such as forest management to reduce fire risks might be taken to reduce the chances of water outages. However, these measures and actions might result in higher water treatment costs and water bills to residential consumers.

In the following sections we ask you to consider which measures and actions you think would be valuable as ways to reduce the chance of water outages.

[INCLUDE IN HOVER LINK for click here for more info line]
(http://www.ecosystemmarketplace.com/pages/dynamic/article.page.php?page_id=7706).
[END OF HOVER LINK]

Water Reliability in Cities and Towns

We are going to present you with different water reliability programs and ask you to choose your preferred program as if you were voting in a referendum. You will vote up to three times and please treat each vote independently for each question. Note that while the questions focus on municipal drinking water management options, industry and other water users would also pay their fair share of any program costs.

We know that how people vote in surveys is often not a reliable indication of how people will actually vote. In surveys some people ignore the sacrifices they would need to make if their vote actually meant they would have less money to spend. In a recent survey like this one, 55% of the people in a community voted for a new program. When the program was put to a vote for real, only 40% actually voted for the program. Therefore, we'd like you to vote in this survey as if your vote was real -- imagine that you actually will have to dig into your pocket and pay the additional charges on your household's water bill if the majority agreed to go ahead with a program.

Some people might choose to vote to keep the current situation because they think:

- It is too much money for the type of benefit I expect to receive.
- The community's tap water supply is reliable enough.
- There are other places where my money would be better spent.

Other people might choose one of the management options because they think:

- The benefits in terms of making water supplies reliable are worth the money.
- This is a good use of money compared to other things I can spend my money on.
- The community tap water isn't very reliable so this would be a good investment.

[PROGRAMMER'S NOTE: The following questions up until the "Follow-up Questions" section can be divided into two main sections with bolded titles: "Managing Your Community's Future Water Supply" and "Managing Alberta's Future Water Supply". Furthermore, within the "Managing Your Community's Future Water Supply" section, there are two subsections entitled "<u>The Future of Water In Your Community: </u>Vote One" and "<u>The Future of Water In Your Community: </u>Vote Two". We would like to randomize the order of the two main sections, as well as randomize the order of Vote One and Vote Two within the "Managing Your Community's Future Water Supply" section. Furthermore, the Vote numbers would change to be consistent with the randomized order including the "<u>The Future of Drinking Water In Alberta: Vote Three" section. For example, for the split sample that faces </u>"<u>The Future of Water In Your Community: </u>Vote Two" first, followed by "<u>The Future of Water In Your Community: </u>Vote One", we want to re-labelled these titles so that the respondent sees Vote One and then Vote Two.]

Managing Your Community's Future Water Supply

We are going to ask you to vote on two different management programs relating to the future of your community's water supply.

<u>The Future of Water In Your Community: Vote One</u>

Suppose that you had a choice between the current situation and a proposed program in a referendum on water reliability.

Current situation: You have indicated that you expect that in a community such as yours there will be about a ____% (transfer responses from Question 24a within program) chance of an unexpected short term (a couple of hours) water outage or reliability problem each year over the next 10 years.

Proposed Management Program: With new investments in management of the water treatment facilities and the watershed, it is estimated that the water outages or reliability problems in your community could be reduced by half to a __________% (divide response from Q24a by 2 and put value in here) chance of a short term water outage or reliability problem each year over the next 10 years.

[OR ALTERNATE OPTION]…could be reduced almost entirely to a less than 1% chance of a short-term water outage or reliability problem each year over the next 10 years.

[PROGRAMMER NOTE: Split sample (randomize) on reduced by "half" and "almost entirely" above and in Question 31]

[PROGRAMMER NOTE: Split sample (randomize) on order of Questions 33, 39 and 44. For Split 1 include all question in their current order. For Split 2, include Question 33 here and the other two Questions (39 and 44) before the valuation questions where indicated later on (Question 39 before Question 37 and Question 44 before Question 42).]

32. If your community holds a referendum to determine whether to put into place the Proposed Management Program and you are asked to vote for or against the program, what would you choose?

Please read the following two statements and choose the one that indicates how you would vote. If you are not currently paying your own water bill, please consider these amounts as increases to your monthly rent (as it is common that rent includes payments for water).

___Yes, I am willing to pay $___ more on my water bill every month ($__ per year) for 10 years starting in January 2015 to pay for the Proposed Management Program that reduces the chance of a short term water outage from ____% (transfer responses from Question 24a within program) by half to a _____________% (divide response from Q24a by 2 and put value in here) chance of an outage each year over the next 10 years.

OR ALTERNATE OPTION…could be reduced almost entirely to a less than 1% chance of a short term water outage or reliability problem each year over the next 10 years.

[PROGRAMMER NOTE: a respondent needs to see half in both parts (or “almost entirely in both).]

___No, I am not willing to pay $___ more on my water bill every month ($__ per year) for 10 years starting in January 2015 to pay for the Proposed Management Program.

33. On a scale from 1 to 5, where 1 is uncertain and 5 is certain, how certain are you that this is the option you would choose if this was an actual vote. Please select one response only.

Uncertain		Somewhat Certain		Certain
1	2	3	4	5

34. To what extent do you believe that the voting results collected from you and other survey respondents will be taken into consideration by policy makers?

Not taken into account				Definitely taken into account
1	2	3	4	5

35. What do you think a person like you in your community (i.e. similar demographics, life situation) would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

36. What do you think the average Albertan would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

When answering this next question, please think about the last vote you completed.

37. Please indicate whether you agree or disagree with the following reasons for why you voted the way you did regarding the Proposed Management Action.

	Strongly Agree	Somewhat Agree	Neither Agree nor Disagree	Somewhat Disagree	Strongly Disagree
I do not believe the program will actually help to make water supplies more reliable.					
I think we should spend whatever it takes to have virtually no water reliability problems.					
I do not believe my individual vote matters in these types of referendum.					
I think this is the best use of my money.					
It is too much money for the benefits.					
The community's tap water supply is sufficient and reliable enough.					
I believe that it is a wise investment that will help prevent water supply problems that might happen in the future.					
I already do things to address my own water reliability problems (e.g. maintain bottled water supplies, have a water filtration system, have rain barrels, etc.).					
I do not trust my community water supplier to ensure water reliability.					
Money spent on these types of projects rarely improves the lives of others.					

The Future of Water In Your Community: Vote Two

Suppose that you had a choice between the current situation and a proposed program in a referendum on water reliability. The proposed program would use one of two methods to improve reliability:

- **Drinking water treatment system:** Investments would be made to upgrade and modernize your community's traditional drinking water treatment system to reduce the annual percentage of water outages occurring in the future. Please click here for examples of specific actions.

[INCLUDED IN HYPERLINK]
Specific drinking water treatment system actions would include:
  - Investing in more modern drinking water treatment systems and increasing the capacity of existing treatment systems,
  - Upgrading and replacing water pipes connecting the water treatment system to households in the community, and
  - Creating more interconnections between drinking water systems and installing backup solutions.

[END OF HYPERLINK]

- **Watershed and forest management**: Investments would be made in the watershed where your drinking water comes from to reduce the potential for events such as forest fires to cause water reliability problems downstream. Please click here for examples of specific activities.

[INCLUDED IN HYPERLINK]
Specific watershed activities would include:
  - Placing buffer strips (i.e. permanent vegetation) along streams to reduce the amount of sediments and debris entering the water,
  - Reducing the amount of hazardous fuels in the watershed to moderate the risk of forest fires, and
  - Forest fire preparedness and response plans to help identify key vulnerabilities and to make responses to fires more effective.

[END OF HYPERLINK]

[PROGRAMMER NOTE: For Split 2, include Question 39 here.]

[PROGRAMMER NOTE: Split sample (randomize) on reduced by "half" and "almost entirely"]

38. Please examine the options below and indicate which option you would vote for.

	Current Situation	Proposed Program
Annual Chance of a Short-term Unexpected Water Outage (a few hours but less than a day)	*(Transfer Responses from Question 24a)* ____________ *%*	Reduced by half to ____________% (divide Q24 a by 2 and put value in here) [OR – if they get the reduce almost entirely…] Reduced almost entirely to a less than 1% chance
Annual Chance of a Longer-term Unexpected Water Outage (at least 1-2 days)	*(Transfer Responses from Question 24b)* ____________*%*	Reduced by half to ____________% (divide Q24 a by 2 and put value in here) [OR – if they get the reduce almost entirely…] Reduced almost entirely to a less than 1% chance
Annual Chance of a Boil Water Advisory	*(Transfer Responses from Question 24c)* ____________*%*	Reduced to by half to ____________% (divide Q24 a by 2 and put value in here) [OR – if they get the reduce almost entirely...] Reduced almost entirely to a less than 1% chance
Method used to improve reliability	Current System	[*Randomize between and include hoverlinks to descriptions]:* Drinking water treatment system improvement Watershed and forest management

Cost of the Program (starting in January 2015)	$0	$___ per year ($__ per month) increase in your water bill for 10 years
Indicate which of the programs above you would vote for if you have to select one of these options.	◻	◻

37A. Please rank the following items. Put a 1 for the item that mattered most to you when you were answering the question, a 2 for the next most important item and a 3 for the item that mattered the least to you.

Rank (1 is mattered most, 3 is mattered least)	
	Annual Chance of a Short-term Unexpected Water Outage (a few hours but less than a day)
	Annual Chance of a Longer-term Unexpected Water Outage (at least 1-2 days)
	Annual Chance of a Boil Water Advisory

39. On a scale from 1 to 5, where 1 is uncertain and 5 is certain, how certain are you that this is the option you would choose if this was an actual vote. Please select one response only.

Uncertain		Somewhat Certain		Certain
1	2	3	4	5

40. To what extent do you believe that the voting results collected from you and other survey respondents will be taken into consideration by policy makers?

Not taken into account				Definitely taken into account
1	2	3	4	5

41. What do you think a person like you in your community (i.e. similar demographics, life situation) would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

42. What do you think the average Albertan would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

Managing Alberta's Future Water Supply

We are now going to ask you to consider province-wide water reliability management programs relating to boil water advisories.

Boil water advisories are issued by Alberta Health Services as preventative measures to protect public health from waterborne infectious agents that may be present in drinking water. If water is consumed without boiling when there is an advisory, serious health problems can arise ranging from moderate illness to, in very rare circumstances, death. For more information on boil water advisories click here.

[INCLUDED IN HYPERLINK]
The three main causes of boil water advisories are

1. High levels of turbidity in the water,
2. Presence of harmful microbes such as E. coli bacteria, and
3. Equipment and process failures or issues.

When a boil water advisory is issued, the public should boil their tap water for drinking, preparing food, beverages, ice cubes, washing fruits and vegetables and brushing teeth. The water should be brought to a rolling boil for 1 minute to kill all disease-causing organisms.

The typical boil water advisory in Alberta last for 9 days.
[END OF HYPERLINK]

Communities have varying chances of being placed under a boil water advisory depending on their source of drinking water and the condition of their water treatment system. The vast majority of boil water advisories are issued for smaller towns and First Nations communities. Approximately 60% of the communities facing boil water advisories in the past have been First Nations communities. The federal government is involved in funding their fair share of water management programs for the First Nations communities under their jurisdiction.

Although the exact numbers change from year to year, over the past 5 years, the average annual number of boil water advisories for different community sizes is presented in the following table:

Community size	Annual number of boil water advisories over past 5 years
Small Communities with less than 500 residents	50
Medium-sized Communities with	4

between 500 and 50,000 residents	
Large Communities with more than 50,000 residents	1

With new investments in management of the water treatment facilities and the watershed, the number of boil water advisories in Alberta could be reduced. Note that given the state of water treatment facilities and the variation in nature (floods, storms, etc.) it may not be possible to completely eliminate boil water advisories.

<u>The Future of Drinking Water In Alberta: Vote Three</u>

Suppose that you had a choice between the current situation and a proposed program in a referendum on water management. The proposed program would be paid through additional income taxes collected on Albertans. The proposed program would use one of two methods to improve reliability:

- **Drinking water treatment system:** Investments would be made to upgrade and modernize traditional drinking water treatment systems across the province to reduce the likelihood of boil water advisories occurring in the future. Please click here for examples of specific actions.

[INCLUDED IN HYPERLINK]

Specific drinking water treatment system actions would include:

- Investing in more modern drinking water treatment systems and increasing the capacity of existing treatment systems,
- Upgrading and replacing water pipes connecting the water treatment system to households in the community, and
- Creating more interconnections between drinking water systems and installing backup solutions.

[END OF HYPERLINK]

- **Watershed and forest management**: Investments would be made in watersheds to reduce the potential for events such as forest fires to cause water reliability problems downstream. Please click here for examples of specific activities.

[INCLUDED IN HYPERLINK]

Specific watershed activities would include:

- Placing buffer strips (i.e. permanent vegetation) along streams to reduce the amount of sediments and debris entering the water,
- Reducing the amount of hazardous fuels in the watershed to moderate the risk of forest fires, and
- Forest fire preparedness and response plans to help identify key vulnerabilities and to make responses to fires more effective.

[END OF HYPERLINK]

[PROGRAMMER NOTE: For Split 2, include Question 44 here.]

43. Please examine the options below and indicate which option you would vote for.

	Current Situation in Alberta	Proposed Program
Annual Number of Boil Water Advisories in **Small Communities with less than 500 residents**	50	(5,15,25,50)
Annual Number of Boil Water Advisories in **Medium-sized Communities with between 500**	4	(1,2,3,4)

and 50,000 residents		
Annual Number of Boil Water Advisories in **Large Communities with more than 50,000 residents**	1	(0,1)
Method used to improve reliability	Current System	Randomize between and include hoverlinks to descriptions: Drinking water treatment system Watershed and forest management
Cost of the Program (starting in 2015)	$0	$___ **per year** (__$ per month) increase in your **provincial income tax** for 10 years
Indicate which of the programs above you would vote for if you have to select one of these options.	□	□

44. On a scale from 1 to 5, where 1 is uncertain and 5 is certain, how certain are you that this is the option you would choose if this was an actual vote. Please select one response only.

Uncertain		Somewhat Certain		Certain
1	2	3	4	5

45. To what extent do you believe that the voting results collected from you and other survey respondents will be taken into consideration by provincial policy makers?

Not taken into account				Definitely taken into account
1	2	3	4	5

46. What do you think a person like you in your community (i.e. similar demographics, life situation) would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

47. What do you think the average Albertan would choose in a referendum like the one just described?
_____ They would vote Yes to the proposed management program
_____ They would vote No to the proposed management program

48. When you think about whether the interests of the population will be taken into account when managing water quality and quantity, to what extent would you trust government resource management institutions?

_____ Completely trust
_____ Somewhat trust
_____ Somewhat not trust
_____ Completely not trust

Follow-up Questions
Now we just have a few more questions to ask you that will help us understand your responses compared to other members of the public.

D1. Did you grow up in a small town or rural area?

___Yes
___No

D2. Do you rent or own the place you currently reside?

___Rent
___Own

D3. Are you a member of a watershed protection community group?

___Yes
___No

D4. Do you consider that the amount of income tax you pay is...? Please select one response only.

___Too high
___About right
___Too low
___Don’t know

D5. Do you consider that the amount you pay for your water bill is...? Please select one response only.

___Too high
___About right
___Too low
___Don’t know

D6. If a provincial election were held today, how would you vote provincially? Please select one response only.

___Alberta Party
___Alberta Liberal Party
___Alberta New Democratic Party
___Progressive Conservative Party of Alberta
___Wildrose Party
___I am not eligible to vote
___I would choose not to vote
___Other (Please type in your response)
___Don’t know
___ Prefer not to say

D7.Compared to others your age, would you say your health is? Please select one response only.

___Much better
___Somewhat better
___About the same
___Somewhat worse
___Much worse
___Don’t know

D8. In the past 12 months, have you ever been a patient overnight in a hospital, nursing home, or convalescent home?

___Yes

___No
___Decline to respond

D9. Which, if any, of the following long-term health conditions do you or members of your family have? Please select all that apply. Please select at least one response (which could be none of the above) in each column.

Health Conditions	Myself	Household Member
Food allergies		
Any other allergies		
Asthma		
Arthritis or rheumatism		
Back problems, excluding arthritis		
High blood pressure		
Migraine headaches		
Chronic bronchitis or emphysema		
Sinusitis		
Diabetes		
Epilepsy		
Heart disease		
Cancer (Please specify type)		
Stomach or intestinal ulcers		
Effects of a stroke		
Any other long-term condition that has been diagnosed by a health professional (Please specify)		
None of the above		

D10. How many individuals live in your household?

______

D11. Are you…?

___Male
___Female

D12. What is your birth date?

______

PROGRAMMER NOTE: Please format D12 so that respondents provide their birthdate by year/month/day.

D13 What is your postal code?

____
___Decline to respond
___Don't know

D13b Could you please provide the first 3 digits of your postal code? We need this information to make sure that survey responses represent the entire province of Alberta.

____
___Decline to respond
___Don't know

E1. Please enter any additional comments you may have about this survey in the space provided.

Thank you for your time. Good bye.