Appalachian State University

Department of Economics Working Paper

Number 20-01 | April 2020

Estimating Discount Rates Using Referendum-style Choice Experiments: An Analysis of Multiple Methods

Gregory Howard
East Carolina University

John C. Whitehead
Appalachian State University

Jacob Hochard
East Carolina University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Estimating Discount Rates Using Referendum-style Choice Experiments: An Analysis of Multiple Methods

Gregory Howard*
East Carolina University

John C. Whitehead
Appalachian State University

Jacob Hochard
East Carolina University

Abstract: There is a growing literature that utilizes stated preference surveys to estimate discount rates. A review of the literature reveals large variation both in the discount rate estimates coming from different stated preference surveys and in the specific methodologies used to estimate discount rates. While most methods use similar theory and logic in deriving discount rate estimates, it is an open question how much of the variation seen in the literature is due to differences in methodology. Using a single data set, we estimate annual discount rates using six different methodologies and find that most of our estimates are tightly clustered between 25-31%. One methodology yields an outlier value of 200%. We also use multiple metrics to examine which methodology yields the "right" discount rate.

Key Words: Discount rate; mixed logit; discrete choice experiment

JEL Codes: D15; Q58; H43; C52

* Corresponding author. Contact information: Email: howardgr@ecu.edu Phone: 252-737-4318
Financial support was provided by East Carolina University's Troxler Endowed Professorship. We are grateful for useful comments provided by seminar participants at Mississippi State University and the 2018 Southern Economics Association Annual Meeting. All remaining errors are our own.

Introduction

Stated preference discrete choice experiments (DCEs) are an analytical tool that can inform preference modelling, especially in areas where revealed preference data are unreliable, incomplete, or scarce. DCEs have been adopted broadly and are common practice in fields from marketing to environmental, health, and transportation economics. These methods are often used to inform public policy decisions, from infrastructure spending to environmental conservation and energy policy. A common feature of DCEs that is often left unexamined is the intertemporal nature of the stream of costs and benefits being posed to respondents. Several approaches have been developed that use respondent decisions to estimate individuals' discount rates. Yet, the ability of these different techniques to deliver consistent discount rate estimates across competing approaches remains an untested source of uncertainty likely to impact policy recommendations.

Typically, respondents are asked to consider paying a one-time or finitely repeated[1] annual tax or fee in exchange for the establishment of some project that delivers a stream of future benefits to the respondent, such as planted forests that improve biodiversity (Yao, et al. 2014), dune restoration that combats coastal erosion (Matthews, et al. 2017) or riparian land restoration that increase recreational amenities (Holland and Johnston 2017). Understanding consumers' intertemporal preferences is crucial to designing optimal policy, as the "best" policy can vary based on the discount rates of the affected population. More patient populations, with lower discount rates, will be more amenable to bearing immediate costs in exchange for more

[1] A minority of studies utilize perpetual annual payments (Egan, et al. 2015), though the assertions that this is a superior method to single or a finite number of payments is the topic of some current debate (Whitehead 2017; Egan, et al. 2018; Whitehead 2018).

distant benefits, while impatient, high discount rate populations are likely to prefer having costs spread over time.

Despite a general tendency in the literature to gloss over intertemporal considerations in DCEs, there is a small but diverse literature that uses stated preference surveys to estimate discount rates. Diverse is an operative word, as these studies vary significantly in their context, in the methodologies they bring to bear when estimating a discount rate, and in the discount rates that they produce. This literature[2] is notable in that there is no dominant technique that has emerged as the standard for estimating discount rates. Rather, multiple techniques have been used, and while these techniques bear the necessary theoretical similarities, they are not identical.

While the use of competing approaches to discount rate estimation is pervasive, authors tend to adopt a single, consistent methodology[3] within a given study. As such, it is not possible to identify whether different estimation methodologies are a source of significant variation in discount rate estimates and the aggressiveness of accompanying policy recommendations. The primary aim of this paper is to compare and examine critically a variety of methodologies used to estimate discount rates in the DCE literature. We adopt six different methods to estimate discount rates for a single data set and find that estimates are, for the most part, relatively consistent, with annual discount rates in the 25-30% range. However, our most sophisticated modelling procedure, which uses a mixed logit estimator with a lognormal distribution for the price coefficient, generates a discount rate of 200%. We also compare the "fit" of elicited

[2] We focus on valuation surveys here. There is another branch of the stated preference literature that uses surveys to directly estimate discount rates (and in many cases risk preferences) without simultaneously valuing a good or service (e.g., Coller and Williams 1999, Benhabib, et al. 2010, Andersen, et al. 2008, Howard 2013).
[3] A small number of studies have used two different methodologies (outlined in Table 1). These studies have generally compared two similar methodologies, and rarely compare less related methodologies such as an endogenous with an ex-post method.

discount rates using the Bayesian Information Criteria (BIC) in addition to an in-sample prediction assessment. Surprisingly, the model with the outlier discount rate possesses the best fit for the data but performs the worst of all methodologies in predictive accuracy. While applied to a different context, this finding mirrors that of Klaiber and Von Haefen (2019), who also document that models using random parameters can improve model fit but decrease predictive accuracy. The variation in our estimated discount rates underscores the need for "Discount Rate Estimation Best Practices" to ensure that the policy recommendations drawn from DCEs are not influenced unnecessarily by major discrepancies among commonly accepted methodologies.

Literature

Table 1 summarizes the literature on estimating discount rates using stated preference valuation surveys. The table highlights several sources of variety in the literature, including the amenity or policy being valued, the structure of the valuation question, the estimation approach used and whether discount rates are derived from delays in benefits or costs to the respondent.

Context: A plurality of studies examine changes in water quality (Stumborg, et al. 2001; Viscusi, et al. 2008; Meyer 2013a; 2013b; Wang and He 2018). The remaining studies cover a variety of topics from ski lift wait times (Crocker and Shogren 1993) to protecting beluga whales (Lew 2018) to reductions in the risk of death (Albertini and Scasny 2011).

Structure of Valuation Question: While early papers tend to use iterative bid intervals or multiple-bid lists in their valuation questions (Crocker and Shogren 1993; Stumborg, et al. 2001; Kovacs and Larson 2008), in the last decade almost all studies have used either multinomial or dichotomous (referendum-style) DCEs (the lone exception to this trend being Wang and He (2018)).

Estimation Approach: There are multiple ways of categorizing estimation approaches. One could compare nonparametric WTP estimators like the Kristrom and Turnbull with parametric models like the probit and logit. Most of the literature utilize parametric methods, with Egan, et al. (2015) and Meyers, et al. (2017) as the primary nonparametric studies.[4] One could further compare procedures that model unobserved heterogeneity, such as mixed logit and generalized multinomial logit models, with simpler models that do not account for this type of heterogeneity. Unobserved heterogeneity models are currently in the minority of studies (Viscusi et al. 2008; Meyer 2013b; Lew 2018), but the literature has seen increased interest in these methods in recent years, a trend that mirrors the wider discrete choice literature. Lastly, one could compare endogenous estimation methods, in which the discount rate is a parameter being estimated in the empirical model, with ex post estimation methods, in which the discount rate is calculated based on model parameters but is not specifically estimated in the model. Prior to 2009, all papers used ex post methods. Bond, et al. (2009) is the first to our knowledge to estimate the discount rate as a parameter in the choice modelling procedure. In the past decade, both ex post and endogenous estimation methods have been utilized with some frequency.

Benefits vs. Costs: In order to estimate a discount rate, the DCE must have between-choice variation in the temporal timing of the effects of the choice. This can be achieved by varying the temporal timing of the benefits (e.g., the policy will be implemented this year vs. 5 years in the future) or the costs (e.g., the policy will be paid for by a fee this year vs. a fee for the next five years). While the early literature was fairly evenly split between discount rates derived from

[4] It is worth noting that these studies also supplement their nonparametric estimates with parametric models as well.

costs and benefits, all post-2013 publications we identified have used cost variations to estimate discount rates.

The substantial heterogeneity that exists in each of these aspects of the literature make it difficult to identify how model-based methodological decisions impact the estimated discount rate. This helps to highlight the need for our approach, where a single data set is used to estimate discount rates using multiple methodologies. This will allow us to identify to what extent methodology influences the resulting discount rate estimate.

Data

Our application is to the control of an invasive species, hemlock woolly adelgid (HWA), in public forests in North Carolina. We fielded the survey from Giguere, Moore and Whitehead (2020, hereafter GMW) with several modifications. In both survey versions, respondents were led through descriptions about ecologically important and socially important areas of hemlock dominated forest. In GMW, there were four levels of acreage for each type of forest: 2500, 5000, 7500 and 10,000. In our application, we include an additional zero level of acreage for both ecological and socially important forest areas in order to test for scope effects at the external and internal margins. To avoid nonsensical policy options, we also included a restriction that omitted any choices with zero values for both types of acreage. Additionally, in GMW, respondents were described biological and chemical treatment methods and the authors found that the biological treatment was preferred. For simplicity, we exclude treatment method as an attribute and present all potential policies as biological treatment.

The prior version of this survey presented respondents with four annual cost amounts, \$50, \$100, \$150 and \$200, which would be paid in each of three years. In our application, we vary the number of years of payment over 3 levels: 1 year, 5 years and 10 years. The annual payment ranges from \$5 to \$1000. The total payment is the product of the annual payment and the payment period and takes on six levels: \$50, \$100, \$250, \$500, \$750, and \$1000.

GMW asked either two or three referendum questions depending on the respondent's survey treatment. In our application, we ask each respondent four referendum questions. As in GMW, respondents were described a referendum: "Imagine that you have the opportunity to vote on the hemlock forest treatment alternative. If more than 50% of North Carolina households vote for the alternative then it would be put into practice." Respondents were then asked, "Would you vote for or against this alternative?" Respondents were given three answer options: I would vote for this alternative, I would vote against this alternative and I don't know how I would vote.

GMW do not use an efficient survey design. In their repeated treatment with three referendum questions, the attributes all vary independently from one another. In contrast, we employ an experimental design with 60 options, which were organized into 15 blocks of 4 each. The efficient design elements, including the total number of choices, attribute levels for each choice, and the specific blocking of the final design, were determined using efficient design macros in SAS including %mktruns, %mktex, %choiceff and %mktblock (Kuhfeld 2003).

In summary, our choice experiment includes four attributes (Table 2). We use two acreage variables (socially and ecologically important), each with 5 levels (0, 2,500, 5,000, 7,500 and 10,000 acres) with a constraint that none of the policies have zeros for both ecologically and socially important acreages. It also has a total payment attribute with 6 levels (50, 100, 250, 500, 750, and 1,000 dollars) and a payment length attribute with 3 levels (1, 5 and 10 years). These two attributes combine to determine the annual payment.

The survey was fielded in December 2017 using an online panel of respondents furnished by Qualtrics. The panel was screened for North Carolina residents with oversampling of western counties. We imposed quotas in gender (50/50 male-female) and education (30% of respondents were either "College graduate or had "Post-graduate training or professional schooling after college").

Table 3 reports a summary of the referendum responses. The percentage of votes in favor of the policy is 40% in the first referendum question, 47% in the second question, 43% in the third question and 40% in the fourth. The annual cost is $154 in the first question and increases in each subsequent question to $208. The average number of years of payment is 5 in the first two questions and 6 in the latter two. The ecologically important acreage is 4234, 6017, 5009 and 4925 in the four referendum questions. The socially important acreage is 4325, 3911, 6781 and 5694 in the four referendum questions.

Models

Each discount rate estimation methodology detailed below shares several important similarities. All methods take as a starting point the assumption of exponential discounting. Each model also assumes additive separability of time periods, either in utility- or WTP-space. Lastly, each model leverages variation in the time period over which households must pay the cost of the policy in order to estimate a discount rate, though the details of how this leverage is gained varies by methodology.

Ex Post Estimation Procedures: Kristrom and Turnbull

Following the literature on nonparametric estimation of WTP (Haab and McConnell 2002; Carson and Hanemann 2005; Whitehead, et al. 1998), we use the proportion of yes votes at different bid levels to estimate a demand curve for the program. We then estimate average WTP as the area under the demand curve. The demand curve is constructed using both the conservative Turnbull approach (Carson and Hanemann 2005), as well as the linear interpolation or Kristrom approach (Kristrom 1990; Boman and Bostedt 1999). Mean WTP is then estimated using the following formulae:

$$\text{Kristrom WTP} = \sum_{j=0}^{J} 0.5 * (\Pr(j) + \Pr(j-1)) * (b_j - b_{j-1}) \quad (1a)$$

$$\text{Turnbull WTP} = \sum_{j=0}^{J} 0.5 * \Pr(j) * (b_j - b_{j-1}), \quad (1b)$$

Where Pr(j) is the proportion of yes votes at bid level j and b_j is bid level j.[5] We use this method to estimate three distinct mean WTP figures: one for each payment horizon (one, five, and ten annual payments). Each mean WTP estimate is then converted to an annual payment (e.g., a total WTP of $500 translates to annual payments of $500, $100, and $50 for one-, five-, and ten-year payment schedules, respectively).

Assuming indifference between WTP estimates for different payment horizons j and k, any two payment schedules yield a unique discount rate by solving

$$\sum_{t=1}^{j} \frac{WTP_j}{(1+d)^{t-1}} = \sum_{t=1}^{k} \frac{WTP_k}{(1+d)^{t-1}} \quad (2)$$

for d. Here, WTP_j is the annual WTP payment for payment horizon j. With three payment horizons, there are three different binary comparisons that can be made between different payment horizons, and as such there is generally no single discount rate that indicates equality

[5] It is additionally assumed that $b_0 = 0$ and Pr(0) = 1, or the proportion of respondents in favor of the program at bid level 0 is 100%. We follow the literature by pooling when the proportion of yes votes is not monotonically decreasing in bid values (Haab and McConnell 2002). With 18 different bids (6 different bid values and 3 different payment horizons), we find nonmonotonicity twice and pool bid proportions appropriately.

between all comparisons. We calculate a discount rate that best fits our data by using these equations to generate differences:

$$Diff_{jk} = \sum_{t=1}^{j} \frac{WTP_j}{(1+d)^{t-1}} - \sum_{t=1}^{k} \frac{WTP_k}{(1+d)^{t-1}}. \tag{3}$$

With three comparisons, we generate three differences for any discount rate. Our estimated discount rate, $d*$, is the value that minimizes the sum of these three squared differences

$$d* = \min_{d} \sum_{t=1}^{3} (Diff_t).^2 \tag{4}$$

Because the Kristrom and Turnbull methods generate different WTP estimates for each payment horizon, each method can generate a different discount rate $d*$.

Ex Post Estimation Procedures: Conditional Logit and Mixed Logit

Both the conditional and mixed logit models, henceforth abbreviated as CL and MXL, are derived from the random utility model where utility from individual n choosing alternative j is comprised of a systematic element, denoted V, and a random error term ε:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \ . \tag{5}$$

Assuming error terms are i.i.d. with a type 1 extreme value distribution, the probability that a respondent will choose alternative j as the best from a set of policy alternatives $\{1,\ldots, J\}$ is given by[6]

[6] As each decision in the data set is a binary referendum-style choice, one could use binary logit or probit models to estimate the following models. We chose to use the multinomial choice framework, as this is a generalized form that is equivalent to the binary logit when there are two alternatives. This framework allows us to use Stata's mixlogit command to model unobserved heterogeneity (Hole 2007).

$$Pr_n(j) = \frac{exp(V_{nj})}{\sum_{k=1}^{J} exp(\mathrm{V}_{nk})}. \tag{6}$$

In the CL model, the systematic element of utility is defined as a function of a vector of non-bid program attributes and alternative-specific constants, **X**, as well as a sequence of cost variables:

$$V_{nj} = \alpha \boldsymbol{X} + \beta_{C1}{*}Cost\_Y1 + \beta_{C5}{*}Cost\_Y5 + \beta_{C10}{*}Cost\_Y10, \tag{7}$$

where *Cost_Yx* is the annual cost of the program to the respondent in year *x*.

The MXL model uses the same set of variables, but further specifies that the vector of coefficients, denoted $\boldsymbol{\beta}$, follows a joint continuous random distribution described by $\boldsymbol{\Omega}$. Each respondent's specific vector of coefficients is a draw from this distribution. Under this assumption, the probability of selecting alternative *j* is now given by

$$Pr_n(j \mid \Omega) = \int_{\boldsymbol{\beta}} Pr_n(j) f(\boldsymbol{\beta} \mid \boldsymbol{\Omega}) d\boldsymbol{\beta}, \tag{8}$$

where $Pr_n(j)$ is given by Equation (6). Each of these models is run using data from all three payment horizons. The coefficient on Year 1 cost, β_{C1}, estimates the present disutility of increasing program cost in Year 1 by $1. Similarly, the coefficient on Year 5 cost, β_{C5}, estimates the present disutility of increasing program cost in Year 5 by $1 while holding Year 1 costs constant. As there is no unique variation in our data set between costs in Years 2-5, β_{C5} is more precisely the present disutility of raising program costs by $1 each year for years 2-5. Following the same logic, β_{C10} captures the present disutility of raising program costs by $1 each year for years 6-10. From this, it follows that the present disutility of increasing the cost of program by $1 per year over 10 years is equal to $\beta_{C1} + \beta_{C5} + \beta_{C10}$.

To estimate a discount rate using these models, we note that the present disutility of increasing the cost of a program by $1 per year over 10 years is equivalent to the disutility of each payment in the year it occurs (β_{C1}), then discounted back to present value. By this logic, we can derive a discount rate by setting these two equivalent terms equal and solving for d:

$$\beta_{C1} + \beta_{C5} + \beta_{C10} = \sum_{t=1}^{10} \frac{\beta_{C1}}{(1+d)^{t-1}}. \tag{9}$$

Endogenous Estimation Procedures: Conditional Logit

When estimating an endogenous discount rate, we amend Equation (7):

$$V_{nj} = \alpha \boldsymbol{X} + \beta_C * PVCost. \tag{10}$$

In this formulation, X is the same vector of non-cost attributes. *PVCost* is defined as the present value of the stream of household payments required by the program:

$$PVCost = \sum_{t=1}^{10} \frac{Cost_t}{(1+d)^{t-1}}.$$

Here, d is a parameter to be estimated along with β_C and the parameter vector α.

Results

Regression Results

Results from our four parametric models, displayed in Table 4, are quite consistent. All models find a negative and statistically significant effect[7] of raising program costs to the household. All

[7] In the mixed logit model with lognormal price distributions, the coefficient is required to have a distribution in only a positive range of values. As such, to compare coefficients from this model to the other three models, one must use the conversion formula -exp(β). This yields values of -0.0082, -0.0030, and -0.0011 for cost in years 1, 5, and 10 respectively. These values are statistically significant (with standard errors using the Delta method, since they are nonlinear combinations of coefficients) for cost in years 1 and 5 (p value < 0.005) but not in year 10 (p value 0.447).

models also produce positive and statistically significant coefficients for both ecologically important variables and only one of the socially important variables (the extensive margin is significant while the intensive margin is not). These results generally mirror the findings of GMW.

Three of the four models also find a positive and statistically significant coefficient for the status-quo alternative-specific constant. The one exception to this trend, the mixed logit with lognormal price distributions, fails to detect a significant effect. Turning to standard deviation estimates for the mixed logit models, there is agreement between models that the status-quo ASC and the dummy variable for any ecologically important acreage have clear preference heterogeneity. Both models fail to detect significant preference heterogeneity for the ecologically important acreage continuous variable, as well as the dummy variable for any socially important acreage. The key difference between models is disagreement regarding the socially important acreage continuous variable. The model that fixes cost preferences finds preference heterogeneity for this acreage attribute, while the model that allows for heterogeneity in cost preferences does not. Lastly, the model that allows for heterogeneity in cost preferences detects heterogeneity for the near-term time period (year 1 and year 5) but fails to detect statistically significant heterogeneity in year 10.

Discount Rate Estimates

Table 5 displays discount rate estimates for each model and method. None of the ex-post estimates possess closed-form solutions, so discount rate standard errors cannot be estimated using model standard errors. As such, standard errors for all discount rate estimates are generated using bootstrapping with 1,000 replications.

The six discount rate estimates can be classified into three groups. First, the ex-post Kristrom and Turnbull estimates yield similar annual discount rate estimates (25-26%). Both are also precisely estimated, with small standard errors, which is unsurprising given that these estimates use similar approaches. A second group of three estimators generate discount rates that are clustered together and are larger than those from the first group. This group, whose estimates range from 30.7% to 31.1%, is more varied in the methodologies employed. It includes both the endogenous and ex-post CL models as well as the ex-post MXL model that assumes a fixed price parameter. Like the first group of estimators, estimates from the second group possess small standard errors. The final group contains a single estimator, the ex-post MXL model that assumes a lognormal distribution for the price coefficient, which appears to be an outlier. This method generates an annual discount rate of 200%. In addition to being substantially larger, this estimate has much larger standard errors (both in an absolute and a relative sense) than the other methodologies. It is likely that this larger, noisier estimate results from the fact that estimated disutility of program cost in year 10, for this model, is very small and not statistically significant.

While one can argue that these groups appear to be different, it is difficult to generate meaningful comparisons of these discount rates, as they are by nature comparisons arising from distinct models (Howard and Liu 2020). We conduct a more rigorous comparison of discount rates by assessing the side-by-side difference in discount rates' point estimates. We then estimate a standard error for this difference using bootstrapping with 1,000 replications, test for a statistical difference between point estimates and repeat the process for all discount rate combinations. With six different discount rates, this generates 15 unique comparisons that are reported in Table 6. We find evidence of differences in estimated discount rates between the MXL model with lognormal cost distribution and all other estimates at the 10% confidence level. No other discount rate estimates appear to differ from each other at any standard level of statistical confidence.

Which Discount Rate is the Right Discount Rate?

It is heartening to see that many different estimation methodologies yield similar discount rates for the same data, but is there a way to identify which discount rate is more accurate when different methodologies generate substantially different estimates? We examine two different methods of evaluating which methodology generates the most appropriate discount rate. One method involves evaluating which underlying model fits the data best. We use the Bayesian Information Criterion (BIC) as our measure of goodness-of-fit. Our second method uses the estimated discount rates, along with other information from the underlying models, to generate model predictions for whether respondents will vote for or against the policy presented to them. Comparing these predictions with observed votes allows us to compare the accuracy or "hit rate[8]" of different models.

Table 7 presents the BIC results and hit rate analyses. One might assume that models that best fit the data will lead to more accurate predictions, and so a reasonable expectation is that BIC and hit rate estimates will identify the same model and discount rate as preferred. Perhaps surprisingly, this is not the case. The evidence suggests that, based on BIC estimates, the more complicated MXL models fit the data best, followed by the ex-post and endogenous conditional logit models. The Kristrom and Turnbull models generate the worst BIC values. In contrast, the

[8] To generate predictions for each choice, we first estimate utility from each alternative as

$$V_{nj} = \alpha \boldsymbol{X} + \beta_{C1} PVCost,$$

where *PVCost* is given by $\sum_{t=1}^{10} \frac{Cost_t}{(1+d)^{t-1}}$. Each alternative is then assigned a probability of being selected using equation (6). The alternative with the highest probability is the predicted choice. Hit rate is then the percentage of predicted choices that match the observed choice.

hit rate for all models are nearly identical (63.69%-63.75%) except for the MXL model with lognormal cost distribution, which achieved a lower hit rate of 61.34%. Thus, the model with the best BIC also appears to possess the worst hit rate. As with the discount rate estimates themselves, we use bootstrapping of the hit rate and BIC differences between methodologies to allow for statistical inference for the inter-methodology comparisons displayed in Tables 8 and 9. All differences between BIC estimates are statistically significant except for those between the Kristrom, Turnbull, and Ex Post CL models.[9] Regarding hit rate, Table 9 confirms that the MXL model with lognormal cost distribution has a significantly lower hit rate than all other models, which are statistically indistinguishable from one another.

Conclusion

Many public policy decisions contain an important intertemporal dimension, with costs and benefits that accrue over different time horizons. Some options may generate substantial initial benefits and are accompanied by costs that occur far into the future. Others may present up-front costs that generate a long temporal stream of benefits. Understanding intertemporal preferences can thus aid policy makers by giving a more holistic understanding of how the public values different policy alternatives. Much of the literature on stated preference valuation has failed historically to account for intertemporal preferences, but in recent years the number of studies that account for intertemporal preferences in stated preference valuation has risen markedly. One facet of this literature is a notable lack of consensus regarding the best methodology to use when

[9] BIC must be the same for the Kristrom and Turnbull estimates, as the underlying models used to estimate them are identical. While these methodologies are nonparametric, their findings can be replicated by estimating a conditional logit model with dummy variables for every combination of total cost to the household and years the cost will be paid over. We use the BIC from this model to compare Kristrom and Turnbull models to the other methodologies.

estimating discount rates. This lack of consensus is in part due to a lack of research that compares different methodologies used in the literature.

This paper is the first to attempt to clarify the ramifications of using different discount rate estimation strategies. Drawing upon a single data set, we generate six different discount rate estimates using a variety of ex post and endogenous estimation methodologies. Reassuringly, we find that most models yield annual discount rates (25-30%) that are reasonable and similar, both to each other and to the broader literature. We do, however, find one methodology that generates a substantially larger discount rate estimate of 200%. This estimate comes from the most sophisticated model analyzed, the mixed logit that allows for unobserved heterogeneity in both price and non-price program attributes. While this model has the best information theoretic statistics (specifically BIC), we find it fares the poorest in terms of prediction accuracy when one uses the estimated discount rate to capture intertemporal preferences.

This research provides two useful lessons moving forward. First, the evidence suggests that many of the potential small changes in how the researcher derives a discount rate from the data have minimal effects on the final estimate (provided each derivation is based on similar assumptions regarding the theory of intertemporal choice). We find similar estimates using ex post and endogenous methodologies, as well as parametric vs. nonparametric methodologies. This is heartening, as it implies that the lack of consensus in the literature regarding the "best" methodology is not inherently problematic, since most competing methodologies are likely to produce similar results. This good feeling, however, should be tempered by the reality that one modeling choice led to a substantially different discount rate.

The second important point is that a better model (defined by goodness-of-fit) does not necessarily produce a better discount rate estimate (i.e., one that more accurately predicts the intertemporal tradeoffs individuals are willing to make). This mirrors the finding of Klaiber and Von Haefen (2019), who see a similar trend where econometrically complex models generate better model fit but worse predictions. Referring to their more complicated models, they note "…these models should generate predictions that are reasonably close if the analyst has correctly specified the underlying data generating process. By implication, the poor in-sample predictions that we find in our empirical applications arise because of model misspecification. This finding represents a cautionary tale to researchers about recent econometric innovations. These models may fail to account for important features of the data that are masked by focusing exclusively on in-sample statistical fit." (Klaiber and Von Haefen 2019, p. 76) This likely explains our MXL model with lognormal price distribution; rather than better representing the underlying data generating process, this modeling framework appears to be overfitting the data, to the detriment of predictive accuracy. This is not to denigrate any specific class of models, but rather to reinforce the truism that complexity should be thoughtfully specified and wielded in the service of specific goals; complexity for complexity's sake has the unfortunate potential to generate more heat than light.

References

Alberini, A. and Ščasný, M., 2011. Context and the VSL: evidence from a stated preference study in Italy and the Czech Republic. *Environmental and Resource Economics*, *49*(4), pp.511-538.

Andersen, S., Harrison, G.W., Lau, M.I. and Rutström, E.E., 2008. Eliciting risk and time preferences. *Econometrica*, *76*(3), pp.583-618.

Benhabib, J., Bisin, A. and Schotter, A., 2010. Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, *69*(2), pp.205-223.

Boman, M. and Bostedt, G., 1999. Valuing the wolf in Sweden: Are benefits contingent on the supply?. In *Topics in environmental economics* (pp. 157-174). Springer, Dordrecht.

Bond, C.A., Cullen, K.G. and Larson, D.M., 2009. Joint estimation of discount rates and willingness to pay for public goods. *Ecological Economics*, *68*(11), pp.2751-2759.

Carson RT, Hanemann WM (2005). "Contingent valuation." In KG M¨aler, JR Vincent (eds.), *Handbook of Environmental Economics*, volume 2, chapter 17, pp. 821–936. Elsevier, New York.

Coller, M. and Williams, M.B., 1999. Eliciting individual discount rates. *Experimental Economics*, *2*(2), pp.107-127.

Crocker, T.D. and Shogren, J.F., 1993. Dynamic inconsistency in valuing environmental goods. *Ecological Economics*, *7*(3), pp.239-254.

Egan, K.J., Corrigan, J.R. and Dwyer, D.F., 2015. Three reasons to use annual payments in contingent valuation surveys: Convergent validity, discount rates, and mental accounting. *Journal of Environmental Economics and Management*, *72*, pp.123-136.

Egan, K.J., Corrigan, J.R. and Dwyer, D.F., 2018. Reply to "a comment on 'three reasons to use annual payments in contingent valuation'". *Journal of Environmental Economics and Management*, *88*, pp.489-495.

Giguere, Christopher, Chris Moore, and John C. Whitehead. "Valuing Hemlock Woolly Adelgid Control in Public Forests: Scope Effects with Attribute Nonattendance." *Land Economics* 96, no. 1 (2020): 25-42.

Haab, T.C. and McConnell, K.E., 2002. *Valuing environmental and natural resources: the econometrics of non-market valuation*. Edward Elgar Publishing.

Hole, A.R., 2007. Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal*, *7*(3), pp.388-401.

Holland, B., and R. Johnston. Optimized quantity-within-distance models of spatial welfare heterogeneity. *Journal of Environmental Economics and Management* 85 (2017): 110-129.

Howard, G., 2013. Discounting for personal and social payments: Patience for others, impatience for ourselves. *Journal of Environmental Economics and Management*, *66*(3), pp.583-597.

Howard, G. and H. Liu. 2020. Empirical blind spot: A note in favor of statistical inference for between-model comparisons. Working Paper.

Klaiber, H.A. and von Haefen, R.H., 2019. Do random coefficients and alternative specific constants improve policy analysis? An empirical investigation of model fit and prediction. *Environmental and Resource Economics*, *73*(1), pp.75-91.

Kovacs, K.F. and Larson, D.M., 2008. Identifying individual discount rates and valuing public open space with stated-preference models. *Land Economics*, *84*(2), pp.209-224.

Kriström, B., 1990. A non-parametric approach to the estimation of welfare measures in discrete response valuation studies. *Land economics*, *66*(2), pp.135-139.

Kuhfeld, W. F. 2003. *Marketing Research Methods in SAS*. SAS Institute Incorporated.

Lew, D.K., 2018. Discounting future payments in stated preference choice experiments. *Resource and Energy Economics*, *54*, pp.150-164.

Matthews, Yvonne, Riccardo Scarpa, and Dan Marsh. "Using virtual environments to improve the realism of choice experiments: A case study about coastal erosion management." *Journal of Environmental Economics and Management* 81 (2017): 193-208.

Meyer, A., 2013a. Estimating discount factors for public and private goods and testing competing discounting hypotheses. *Journal of Risk and Uncertainty*, *46*(2), pp.133-173.

Meyer, A., 2013b. Intertemporal valuation of river restoration. *Environmental and Resource Economics*, *54*(1), pp.41-61.

Myers, K., Parsons, G. and Train, K., 2017. Inadequate response to frequency of payments in contingent valuation of environmental goods. *Contingent Valuation of Environmental Goods, Edward Elgar Publishing, Inc., Massachusetts, USA*, pp.43-57.

Stumborg, B.E., Baerenklau, K.A. and Bishop, R.C., 2001. Nonpoint source pollution and present values: A contingent valuation study of Lake Mendota. *Review of Agricultural Economics*, *23*(1), pp.120-132.

Vasquez-Lavín, F., Oliva, R.D.P., Hernández, J.I., Gelcich, S., Carrasco, M. and Quiroga, M., 2019. Exploring dual discount rates for ecosystem services: Evidence from a marine protected area network. *Resource and Energy Economics*, *55*, pp.63-80.

Viscusi, W.K., Huber, J. and Bell, J., 2008. Estimating discount rates for environmental quality from utility-based choice experiments. *Journal of risk and Uncertainty*, *37*(2-3), pp.199-220.

Wang, H. and He, J., 2018. Implicit individual discount rate in China: A contingent valuation study. *Journal of environmental management*, *210*, pp.51-70.

Whitehead, J.C., 2017. Who Knows What Willingness to Pay Lurks in the Hearts of Men? A Rejoinder to Egan, Corrigan, and Dwyer. *Econ Journal Watch*, *14*(3), p.346.

Whitehead, J.C., 2018. A comment on “Three reasons to use annual payments in contingent valuation”. *Journal of Environmental Economics and Management*, *88*, pp.486-488.

Whitehead, J.C., Haab, T.C. and Huang, J.C., 1998. Part-whole bias in contingent valuation: Will scope effects be detected with inexpensive survey methods?. *Southern Economic Journal*, pp.160-168.

Yao, Richard T., Riccardo Scarpa, James A. Turner, Tim D. Barnard, John M. Rose, João HN Palma, and Duncan R. Harrison. "Valuing biodiversity enhancement in New Zealand's planted forests: socioeconomic and spatial determinants of willingness-to-pay." *Ecological Economics* 98 (2014): 90-101.

Table 1: Literature on Discount Rates in Stated Preference Choice Experiments

Paper	Context	Method	Discount Rate	Delay	Survey Instrument
Crocker and Shogren 1993	Reducing ski lift wait times	Ex post	0.041-0.368 (per minute)	Benefits	Stated WTP using iterative bid intervals
Stumborg, et al. 2001	Improving water quality in Lake Mendota	Ex Post Tobit	0.40	Costs	Multiple Price List Table from $0 to $300
Kovacs and Larson 2008	Expanding a public park	Ex Post Probit	0144-.369	Costs	Referendum-style DCE with iterative bids
Viscusi et al 2008	Improvements in water quality	Ex Post Conditional Logit	0.61-0.77	Benefits	DCE
Viscusi et al 2008	Improvements in water quality	Ex Post Mixed Logit (not correlated distributions)	0.63-0.79	Benefits	DCE
Bond et al. 2009	Protection of sea lion critical habitat areas	Endogenous, Probit	0.231	Costs	Referendum-style DCE
Albertini and Scasny 2011	Reductions in risk of death	Endogenous, Conditional Logit	0	Benefits	DCE
Meyer 2013a	Improvements in water quality	Endogenous, Conditional Logit	0.10-0.11	Benefits	DCE
Meyer 2013b	Improvements in water quality	Endogenous, Mixed Logit	0.11-0.12	Benefits	DCE

Table 1 Continued: Literature on Discount Rates in Stated Preference Choice Experiments

Paper	Context	Method	Discount Rate	Delay	Survey Instrument
Egan, et al. 2015	Watershed restoration	Ex Post Turnbull	0.15-1.04	Costs	Referendum-style DCE
Meyers, et al. 2017	Protecting migratory shorebirds	Ex Post Turnbull	8.37	Costs	Referendum-style DCE
Meyers, et al. 2017	Protecting migratory shorebirds	Ex Post Probit	3.51	Costs	Referendum-style DCE
Vasquez-Lavin, et al. 2019	Funding Marine Protected Areas	Endogenous Probit	0.70-3.72	Costs	5 Referendum-style DCEs
Lew 2018	Protecting beluga whales	Endogenous, Mixed Logit	1.57	Costs	DCE
Lew 2018	Protecting beluga whales	Endogenous, Generalized MNL	1.28	Costs	DCE
Wang and He 2018	Water quality and landscape improvements	Ex Post Probit	1.41-3.15	Costs	Multiple Bound Dichotomous Choice

Table 2. Choice Experiment Referendum Question, Attributes and Levels

<table>
<tr><td colspan="2">Imagine that you have the opportunity to vote on the hemlock forest treatment alternative. If more than 50% of North Carolina households vote for the alternative then it would be put into practice.

Consider the following alternative:

The area treated of ecologically important hemlock dominated forests is _____ acres, _____% of the total. The area treated of socially important hemlock dominated forests is _____ acres, _____% of the total.

(1 year version) The cost to you is $_____ next year. You would pay $_____ total.
(5 or 10 year versions) The cost to you is $_____ annually for the next _____ years. You would pay $_____ total.

Would you vote for or against this alternative?

I would vote for this alternative
I would vote against this alternative
I don’t know how I would vote</td></tr>
<tr><td>Attribute</td><td></td></tr>
<tr><td>Ecologically important acreage</td><td>5 levels: 0 acres, 0% of total
2,500 acres, 8% of total
5,000 acres, 17% of total
7,500 acres, 25% of total
10,000 acres, 33% of total</td></tr>
<tr><td>Socially important acreage</td><td>5 levels: 0 acres, 0% of total
2,500 acres, 8% of total
5,000 acres, 17% of total
7,500 acres, 25% of total
10,000 acres, 33% of total</td></tr>
<tr><td>Payment Years</td><td>3 levels: 1 year
5 years
10 years</td></tr>
<tr><td>Total Cost</td><td>6 levels: $50
$100
$250
$500
$750
$1,000</td></tr>
</table>

Table 3. Referendum Data Summary

Question	Vote	Annual Cost	Years	Ecological	Social
1	40.07%	135.23	4.96	4234	4325
2	46.73%	153.85	5.26	6017	3911
3	42.67%	191.39	6.25	5009	6781
4	40.34%	208.33	6.31	4925	5694
Sample size = 2266					

Table 4. Logit Regression Results

	Conditional Logit		Mixed Logit – fixed cost		Mixed Logit – lognormal cost		Conditional Logit – Endogenous DR	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Cost Year 1	-0.0023***	0.0001	-0.0036***	0.0002	-4.8008***	0.0944	-	-
Cost Year 5	-0.0050***	0.0004	-0.0082***	0.0006	-5.7955***	0.3269	-	-
Cost Year 10	-0.0018***	0.0007	-0.0025**	0.0011	-6.8344***	1.3149	-	-
Eco Acreage	0.0020**	0.0008	0.0033***	0.0012	0.0058***	0.0014	0.0020**	0.0008
Any Eco Acreage	0.5126***	0.0730	0.7194***	0.1107	0.7714***	0.1226	0.5112***	0.0725
Soc Acreage	-0.0010	0.0008	-0.0019	0.0012	-0.0011	0.0013	-0.0010	0.0008
Any Soc Acreage	0.1776**	0.0699	0.2987***	0.1060	0.4167***	0.1186	0.1761**	0.0697
Status Quo ASC	0.2947***	0.0791	0.3756***	0.1143	0.0185	0.1279	0.2946***	0.0791
PV cost	-	-	-	-	-	-	-0.0023***	0.0001
Discount Rate	-	-	-	-	-	-	0.3071***	0.0377
SD Estimates								
Eco Acreage	-	-	0.0051	0.0044	0.0026	0.0036	-	-
Any Eco Acreage	-	-	0.6561***	0.1878	0.7984***	0.1931	-	-
Soc Acreage	-	-	0.0086***	0.0026	0.0054	0.0037	-	-
Any Soc Acreage	-	-	0.0788	0.2190	0.0773	0.2687	-	-
Status Quo ASC	-	-	1.5983***	0.0737	1.2954***	0.1225	-	-
Cost Year 1	-	-	-	-	1.9756***	0.1272	-	-
Cost Year 5	-	-	-	-	1.0761***	0.3996	-	-
Cost Year 10	-	-	-	-	1.4089	1.0527	-	-
LL	-5761.2		-5290.0		-5163.6		-5761.2	
Individuals	2,266		2,266		2,266		2,266	
Sample	18,128		18,128		18,128		18,128	

Notes: *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. Lognormal price coefficients are translated to standard cost coefficients by taking the negative exponent.

Table 5: Discount Rate Estimates by Methodology

Method		Discount Rate	Std. Err.
Ex Post Kristrom		0.260***	0.031
Ex Post Turnbull		0.250***	0.086
Ex Post CL		0.311***	0.045
Endogenous CL		0.307***	0.040
Ex Post MXL	Fixed Cost	0.309***	0.040
Ex Post MXL	LogN Cost	2.00**	1.016

Notes: *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. Standard errors are estimated using bootstrapping with 1,000 replications.

Table 6: Test for Differences in Discount Rate Estimates by Methodology

	Kristrom	**Turnbull**	**Ex Post CL**	**Endogenous CL**	**Ex Post MXL Fixed**	**Ex Post MXL LogN**
Kristrom						
Turnbull	0.895					
Ex Post CL	0.288	0.473				
End CL	0.269	0.489	0.865			
MXL Fixed	0.262	0.482	0.929	0.941		
MXL LogN	**0.086**	**0.084**	**0.093**	**0.091**	**0.092**	

Notes: Reported values are *P* values for a test of equality between the two methodologies. Bolded values indicated *P* values of 0.1 or less. Standard errors for differences are calculated using bootstrapping with 1,000 replications.

Table 7: Goodness-of-Fit Measures by Methodology

Method	Discount Rate	Std. Err.	BIC	Std. Err.	Prediction Hit Rate	Std. Err.
Ex Post Kristrom	0.260***	0.031	11,666.49***	103.83	0.6375***	0.005
Ex Post Turnbull	0.250***	0.086	11,666.49***	103.83	0.6375***	0.005
Ex Post CL	0.311***	0.045	11,600.91***	70.80	0.6369***	0.005
Endogenous CL	0.307***	0.040	11,591.13***	70.81	0.6369***	0.005
Ex Post MXL, Fixed Price	0.309***	0.040	10,707.49***	83.93	0.6369***	0.006
Ex Post MXL, Lognormal Price	2.00**	1.016	10,484.08***	88.87	0.6134***	0.007

Notes: Standard errors presented in parentheses. * and **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. Standard errors are estimated using bootstrapping with 1,000 replications.

Table 8: Test for Differences in BIC Estimates by Methodology

	Kristrom	**Turnbull**	**Ex Post CL**	**Endogenous CL**	**Ex Post MXL Fixed**	**Ex Post MXL LogN**
Kristrom						
Turnbull	1.000					
Ex Post CL	0.293	0.293				
End CL	0.227	0.227	**< 0.005**			
MXL Fixed	**< 0.005**	**< 0.005**	**< 0.005**	**< 0.005**		
MXL LogN	**< 0.005**	**< 0.005**	**< 0.005**	**< 0.005**	**< 0.005**	

Notes: Reported values are *P* values for a test of equality between the two methodologies. Bolded values indicated *P* values of 0.05 or less. Standard errors for differences are calculated using bootstrapping with 1,000 replications.

Table 9: Test for Differences in Hit Rate Estimates by Methodology

	Kristrom	**Turnbull**	**Ex Post CL**	**Endogenous CL**	**Ex Post MXL Fixed**	**Ex Post MXL LogN**
Kristrom						
Turnbull	1.000					
Ex Post CL	0.889	0.889				
End CL	0.890	0.890	1.000			
MXL Fixed	0.891	0.891	1.000	1.000		
MXL LogN	**< 0.005**	**< 0.005**	**< 0.005**	**< 0.005**	**< 0.005**	

Notes: Reported values are *P* values for a test of equality between the two methodologies. Bolded values indicated *P* values of 0.05 or less. Standard errors for differences are calculated using bootstrapping with 1,000 replications.