

# Sensitivity of WTP Estimates to Definition of ‘Yes’: Reinterpreting Expressed Response Intensity

Mimako Kobayashi, Kimberly Rollins, and M. D. R. Evans

Willingness to pay (WTP) estimation typically involves some strategy for mapping non-dichotomous contingent valuation (CV) responses onto a dichotomous yes/no dependent variable. We propose a new approach to selecting which responses qualify as ‘yes.’ We apply the proposed method to polychotomous CV data for preventative land management programs in the Great Basin. We also estimate WTP using other methods of response recoding found in the literature. By contrasting the results under different approaches, we demonstrate how and why WTP point estimates vary across recoding methods and discuss the comparative advantages of our more generalized recoding approach that is based on predicted probabilities of ‘yes’ responses.

**Key Words:** contingent valuation, Great Basin, invasive weeds, land management, polychotomous format, response intensity, response mapping, wildfire

It is especially important that researchers be cautious about presenting willingness to pay (WTP) point estimates to policy makers when the estimates are sensitive to subjective decisions that the researchers make during the estimation process. It is well known that WTP point estimates are sensitive to the choice regarding which observations should be included in a WTP estimation model as ‘yes’ (or “willing”) responses and which as ‘no’ (“not willing”) responses, when contingent valuation (CV) data are collected in formats that accommodate varying levels of “response intensity”<sup>1</sup> (Akter et al. 2009, Champ, Alberini, and Correas 2005, Shaikh, Sun, and van Kooten 2007, Vossler et al. 2004, Samnaliev, Stevens, and More 2006, Broberg and Brannlund 2008). However,

because there is no standardized procedure for mapping non-dichotomous CV responses into the yes/no binary variable, it is the responsibility of the researcher to justify an appropriate recoding rule.

With this paper, we wish to stimulate discussion by proposing a new approach to selecting which responses qualify as ‘yes’ that (1) uses more information contained in a data set than do others in the literature, (2) does not rely on the assumption that response intensity is solely a function of respondent uncertainty about preferences, and (3) is applicable regardless of the underlying sources of variation in response intensity. The need for the discussion stems from the fact that estimation of WTP using random utility models (RUMs) typically requires responses to be coded as ‘yes’ or ‘no’ for a given bid amount even when CV data are not collected in a dichotomous choice format. In fact, since the National Oceanographic and Atmospheric Administration (NOAA) Panel on Contingent Evaluation recommended

---

Mimako Kobayashi is Research Assistant Professor and Kimberly Rollins is Associate Professor in the Department of Resource Economics, and M. D. R. Evans is Associate Professor in the Department of Resource Economics and the Department of Sociology, at the University of Nevada, Reno, in Reno, Nevada.

This is Contribution Number 33 of the Sagebrush Steppe Treatment Evaluation Project (SageSTEP), funded by the U.S. Joint Fire Science Program. Support was also provided by the Nevada Agricultural Experiment Station. The authors wish to acknowledge comments and suggestions received from members of the W2133 Project at their 2009 annual meeting in Austin, Texas, and from two anonymous reviewers.

---

<sup>1</sup> Response direction refers to the respondent’s ‘yes’ or ‘no’ attitude, while response intensity refers to the strength of this attitude.

that questionnaires used to collect CV data should include an alternative category to a 'yes' or 'no' response (NOAA 1993),<sup>2</sup> polychotomous and other formats that allow finer measurements of response intensity have become standard practice.<sup>3</sup> As a result, a WTP estimation stage typically involves some strategy for mapping CV responses that include various levels of response intensity onto a dichotomous dependent variable. This process is subject to error due to imprecision in interpreting expressed response intensity because the researcher does not know the true underlying functional relationship that maps intensity of responses onto the binary variable.

Individuals can express response intensity differently, which necessitates the response recoding strategy also to vary across individuals. Reasons for varied response intensity include preference uncertainty (Alberini, Boyle, and Welsh 2003, Ready, Whitehead, and Blomquist 1995); ambiguity arising from context uncertainty about how a policy change might affect utility (Ariely, Loewenstein, and Prelec 2006, Cameron and Englin 1997); inability or unwillingness to exert the effort required for full consideration of the choice task (Krosnick 1991, Ready, Whitehead, and Blomquist 1995); differences in individual-specific tendencies toward expressing intensity of attitude (Hui and Triandis 1985, Hamilton 1968, Albaum et al. 2007); and random measurement error due to instrument design (DeShazo and Fermo 2002, DeShazo 2002, Groves 1991, Vossler et al. 2004, Bishop 1987). The method presented in this paper does not intend to identify or differentiate reasons for variation in expressed response intensity. Rather, it accommodates any reasons while controlling for individual differences in expressing response intensity.

We apply the proposed method to polychotomous data collected in a pilot study, to generate valuation estimates for preventative land management programs in the Great Basin that would arrest the accelerated wildfire cycle due to invasive annual weeds. We also estimate WTP using other methods of response recoding found in the literature. By contrasting the results, we demonstrate how and why WTP point estimates vary across recoding methods and discuss the problems of adopting an ad hoc strategy.

The organization of this paper is as follows. The next section discusses how response intensity is treated in the WTP literature, and introduces a method for coding responses into a dichotomous dependent variable for use in estimating WTP that addresses problems found in previous studies. The third section describes the data used to apply our method. The fourth section presents estimation models, with the subsequent section discussing results. The last section concludes the paper and presents directions for future research.

## Response Intensity and Response Recoding

In the WTP literature, the error associated with mapping non-dichotomous CV responses onto a dichotomous variable to accommodate WTP estimation has not explicitly been recognized; rather, the task has been treated as a practical issue of recoding. Where a CV format includes a third response category such as 'don't know' (DK) or 'not sure' (NS) for a given bid amount in addition to 'yes' and 'no,' typical recoding rules are either to drop DK responses or to recode them as either 'yes' or 'no' to generate a dichotomous dependent variable (Champ, Alberini, and Correas 2005). In a polychotomous choice format,<sup>4</sup> where

<sup>2</sup> The original intention of the recommendation to include a third category, generally 'don't know' (DK) or 'not sure' (NS), was 1) to reduce potential upward biases in WTP estimates originating from the hypothetical nature of referenda format that could lead to 'yes' responses to bid amounts above an individual's true maximum WTP (hypothetical bias; Champ and Bishop 2001); and 2) to reduce item nonresponses by respondents who genuinely hold opinions by providing a better representation of responses that are not genuinely 'yes' or 'no.'

<sup>3</sup> The three most common CV formats are 1) to include a third category, generally 'don't know' or 'not sure'; 2) a polychotomous format that includes, for example, 'definitely yes,' 'probably yes,' 'probably no,' and 'definitely no'; and 3) a dichotomous choice question with a follow-up question asking the respondent to indicate how certain they are of their response on a scale of 1 to 10.

<sup>4</sup> The polychotomous choice (PC) format is widely used throughout the social sciences because it allows respondents to represent their underlying attitudes in a way that is intuitively appealing to them (Tourangeau, Rips, and Rasinski 2000, Schwartz and Sudman 1995), and it is robust to replications with test-retest reliabilities tending to be high (Krosnick, Judd, and Wittenbrink 2005). First introduced to the valuation literature by Ready, Whitehead, and Blomquist (1995), the PC format presents multiple response options to the bid amount, which vary by both direction (yes/no) and intensity (including level of certainty). Whitehead et al. (1998) conduct a series of validity tests for dichotomous and polychotomous choice data and conclude that the PC format produces theoretically valid WTP estimates that are convergently valid and that the PC format is a valuable method especially when intensity of respondent preferences are an important consideration.

the level of response intensity is directly incorporated into the response options—e.g., 'definitely yes' (DY); 'probably yes' (PY); 'probably no' (PN); 'definitely no' (DN); and DK/NS—typical recoding strategies are to use DY or DY + PY as 'yes' and the rest as 'no' while treating DK/NS responses similarly to the previous case (Ready, Whitehead, and Blomquist 1995, Welsh and Poe 1998, Whitehead et al. 1998, Alberini, Boyle, and Welsh 2003). Another CV format, first introduced by Li and Mattsson (1995), uses the individual's subjective probability that their valuation is greater than the bid (for a 'yes') or less than the bid (for a 'no'), obtained through a follow-up question where respondents are asked to rate how certain they are about their dichotomous choice on a 10- or 100-point scale (often referred to as the numerical certainty scale or NCS).<sup>5</sup> For this CV format, 'yes' and 'no' responses are typically recoded according to the researcher's chosen level of certainty; e.g., if the subjective probability of a 'yes' response is below 20 percent, recode it as 'no' (e.g., Champ et al. 1997, Loomis and Ekstrand 1998, Samnaliev, Stevens, and More 2006, Shaikh, Sun, and van Kooten 2007, Akter et al. 2009).<sup>6</sup>

These response coding rules adopted in the WTP literature, however, are typically applied uniformly to all observations in a data set, thereby dismissing variation across individuals in the way of expressing response intensity, which is consid-

ered to contribute to random measurement error (Bollen 1989, Kelley 1973, Li and Mattsson 1995).<sup>7</sup> In the valuation literature, response intensity has for the most part been considered as an expression of the certainty with which respondents answer the question at hand; that is, their certainty over whether or not they would be willing to pay the presented bid amount in order to achieve the proposed policy change. In the literature from other social sciences on survey format and response, intensity is an expression of certainty but also of several other response tendencies (Tourangeau, Rips, and Rasinski 2000, Krosnick, Judd, and Wittenbrink 2005). As a result, individual-level differences in response tendencies could lead two or more individuals having the same underlying latent value for WTP to express response intensity differently (e.g., 'definitely yes' vs. 'probably yes'). Given that reasons for expressing different response intensities can range widely and be specific to the individual, application of a uniform recoding rule to all individuals in a data set is not sensible.

Even when considering how only respondent uncertainty can affect response intensity, there is not a singular notion of what is meant by respondent uncertainty or how it affects expressions of response intensity in various CV question formats. Responses of less than the strongest intensity, such as DK, PY, and PN, can be motivated by a variety of reasons (Groothuis and Whitehead 2002, Wang 1997, Haener and Adamowicz 1998, Caudill and Groothuis 2005, Champ, Alberini, and Correas 2005, Mozumder and Berrens 2007, Blumenschein et al. 2008). Wang (1997) argues that an individual chooses DK when the presented bid is very close to his or her actual WTP and that exclusion of DK responses implies tossing out data of potentially great informational content. Broberg and Brannlund (2008) consider "imprecision" in preferences where people wish to state WTP in intervals rather than precise values, and the width of the intervals increase with preference uncertainty. DK responses can also arise due to

<sup>5</sup> While the method continues to be widely used in the applied literature (see for example, Samnaliev, Stevens, and More 2006, Shaikh, Sun, and van Kooten 2007, Akter et al. 2009), it suffers from several problems. First, given that individuals are assumed to be uncertain about preferences, it is inconsistent to assume they are sure about the level of certainty in their response. Second, subjective probabilities of response certainty do not account for variation in individual-specific response tendencies across the sample. Third, while it may appear to the researcher that the fine levels of gradation provided in formats using numerical certainty scales provide a cardinal measure of uncertainty, the categories listing "probability of certainty" may still be regarded as ordinal measures and subject to random measurement error. Ready, Navrud, and Dubourg (2001) nicely sum this up: "To respondents, 95% sure may mean the same thing as 90%, or 99%, or even 100% sure. Finer gradations may not generate more information if respondents are using only three or five rough categories of certainty." Fourth, this format is known to have lower test-retest reliability and generates excess missing data (Krosnick 1999). At any rate, the need for response recoding still arises and the methods that we describe in this paper would similarly apply.

<sup>6</sup> Exceptions are Li and Mattsson (1995) and Aadland and Caplan (2006). While they take yes/no responses at face value, they use the information collected through the certainty scale to adjust the variance estimates (Li and Mattson 1995) and as an explanatory variable in estimating WTP (Aadland and Caplan 2006).

<sup>7</sup> Whether captured in a DK/NS response category, in various response intensity categories (definitely, probably, etc.), or in a numerical certainty scale, response intensities expressed by individuals are inherently subject to random measurement error (Groves 1991, Fuller 1986, Carroll et al. 2006).

uncertainty on the part of respondents about the level of benefits they personally will receive from a public good (Mozumder and Berrens 2007, Alberini, Boyle, and Welsh 2003). Such preference uncertainty may originate from poor definition in survey instruments of the good in question, respondents' lack of experience with the good, missing information about the good, or state-dependent circumstances under which the good will be available to them. Respondents may also have internal conflicts in evaluating the benefits of the good. Bateman, Burgess, and Hutchinson (2008) argue that preference uncertainty may constitute a source of bias that overwhelms other sources of bias, including hypothetical bias, and that this area needs more research. Thus, it would seem that individual stated expressions of intensity (whether via polychotomous response categories or via a numerical certainty scale) would be subject to measurement error.

An issue less-frequently discussed in the valuation literature is individual tendencies to choose extreme responses or to stay away from extremes (Greenleaf 1992). Response tendencies are related to individual motivation and certainty (Cantril 1946); anxiety level (Lewis and Taylor 1955, Crandall 1965); and personality characteristics (Hamilton 1968). Elliot et al. (2009) find race/ethnicity and socio-economic status are associated with extreme responses, and Naemi, Beal, and Payne (2009) find personality traits, specifically intolerance of ambiguity, simplistic thinking, and decisiveness, are associated with overuse of endpoints of scale-type questions. When dealing with cognitively difficult questions, respondents may shift to a satisficing response strategy where attempts are made to meet criteria for "adequacy" rather than expending effort on generating "optimal" answers (Krosnick 1999). In these cases, individuals may be more likely to respond with a "safe" answer, choosing more neutral points on a rating scale, endorsing the status quo, or choosing DK to avoid expending effort and taking possibly risky stands. It is also found that an individual's response tendencies can shift during a survey (Hui and Triandis 1985). In the case of polychotomous choice format, interpretation of statements such as 'probably' and 'maybe' likely varies by individuals (Broberg and Brannlund 2008).

It is likely that any given sample includes responses that fit several of these reasons for variation in stated response intensity. Accordingly, much of the literature focuses on how to handle "middle" responses, or responses that are not in the categories with strongest intensity, in estimating WTP when the underlying respondent motivations for choosing those categories are unknown. The two common approaches to recode middle responses are to impose rules on an entire category (e.g., all DK responses are dropped or coded as 'no')<sup>8</sup> and to apply rules to individual responses (e.g., some DK responses are coded as 'no' but others as 'yes'). Among authors who apply the latter recoding strategies that vary depending on the individual, Haener and Adamowicz (1998) use responses to a follow-up question to recode DK responses into either 'yes' or 'no,' and find that the goodness of fit in WTP estimation is improved over the alternate models with dropped DK and with DK recoded as 'no.' Alternatively, Caudill and Groothuis (2005) estimate the probabilities of 'yes,' 'no,' and DK for each of the DK observations using multinomial logit combined with an EM (expectations maximization) algorithm for missing data problems. They subsequently recode the original DK responses to the categories with the highest predicted probabilities. The methods by Haener and Adamowicz (1998) and Caudill and Groothuis (2005) each produce a single recoded data set for WTP estimation. In this paper we propose an alternative method where a series of data sets is generated, with which a series of WTP estimates is obtained.

The response mapping strategy proposed in this paper is the following. First, we use expressed intensity response as the dependent variable in a multinomial logit model to estimate choice probabilities. In our application, we have five response categories: DY, PY, PN, DN, and NS, where DY indicates the strongest intensity of a 'yes.' We use the resulting coefficient estimates to calculate the predicted probability of each individual's

<sup>8</sup> Several authors statistically investigate whether DK responses collectively are different from 'yes' or 'no' responses. Some studies find that DK/NS responses would have been 'no' (Carson et al. 1998; WTP format in Groothuis and Whitehead 2002), while others find that DK responses are neither 'yes' nor 'no' but appear to be a third category altogether (Groothuis and Whitehead 2002, Haener and Adamowicz 1998).

response being in the strongest 'yes' intensity category (DY), or  $p^{\wedge}(DY)$ . Then we code responses as 'yes' if  $p^{\wedge}(DY)$  is greater than specified cutoff levels (0.9, 0.8, 0.7, etc.) and code them as 'no' otherwise. This strategy results in a separate data set for each cutoff point, or  $p^{\wedge}(DY)$  level. WTP estimation is applied to each of these data sets, thus generating a series of WTP estimates.<sup>9</sup> Each WTP estimate reflects an observation inclusion criterion that uses an explicit minimum level of predicted probability of being in the highest intensity category for a 'yes' response. Thus, confidence intervals for each WTP estimate are also based on the specific inclusion criterion for the observations in that data set.

By using individual variation in data to predict DY probabilities, we take advantage of the information contained in the data and avoid the need to arbitrarily impose a uniform rule on the entire data set. The strategy accommodates uncertainty as well as other reasons for expressed response intensity by including in the probability estimation model variables that characterize individuals and their attitudes towards the good in question. In particular, this approach is capable of controlling for possible effects of individual response tendencies and variability in subjective interpretation of response categories. While probabilities collected using numerical certainty scale formats (e.g., Li and Mattsson 1995) are individuals' *stated* probabilities of certainty about their yes/no response, we use stated responses to estimate probabilities of each of the response categories. The predicted probability of a DY response from the first-stage multinomial logit takes into account the probabilities of each respondent choosing any of the other intensity categories.

The strategy by Caudill and Groothuis (2005) is the closest to ours in motivation and the method employed. However, as it turns out, their approach predetermines the observation inclusion criterion according to the number of response categories included in the CV questionnaire. In their application, they use a multinomial logit to estimate for each DK observation the predicted

probabilities of the three response categories ('yes,' 'no,' and DK). They recode DK observations into the category with the highest predicted probability. This means that a DK response could be recoded to a 'yes' regardless of the magnitude of predicted probability of 'yes.' With three response categories, this implies that an observation with a predicted probability of 'yes' as low as 34 percent can be coded as 'yes' (predicted probabilities in the other two categories each should be up to 33 percent). If five categories had been used, the inclusion criterion would recode to 'yes' responses with up to a 21 percent chance of being in the strongest 'yes' intensity category (the other four categories would each be less than 20 percent). As the number of intensity categories increases, using their method, the inclusion criterion (selecting the observations where the highest probability is for the highest intensity 'yes' response) results in recoding of more observations that are less likely to be 'yes' as 'yes' responses. Therefore, larger numbers of response categories likely results in higher WTP estimates. Instead of arbitrarily choosing inclusion criterion, our method successively changes the predicted probability cutoff points, so that a range of WTP estimates is generated, rather than only one that is subject to the researcher's arbitrary choice of cutoff point.

We apply this method to data collected using a polychotomous CV format in a pilot study, to generate valuation estimates for preventative land management programs in the Great Basin that would arrest the accelerated wildfire cycle due to invasive annual weeds. Note that the approach is also appropriate for data collected using a dichotomous format with follow-up certainty scale questions.

## Survey Design and Data

The Great Basin region of the western United States has experienced rapid population growth and development over the last thirty years. Wildfire has always been a natural part of the landscape in the region; however a trend towards much more frequent and severe wildfires has occurred due to the rapid increase in invasive annual grasses. These grasses are highly flammable, they outcompete native grasses and shrubs,

<sup>9</sup> While our data were generated using a polychotomous question format where attitude direction and intensity are included in a single question, the same strategy could be applied to a dichotomous choice question with a follow-up second question that measures intensity. The multinomial logit stage would include more choice options (for example 20, if a 10-point scale is used in the follow-up question).

and once they become the dominant species in an area, they provide the fuel that leaves the area vulnerable to wildfires every three to five years, instead of the fifty years to seventy years that is standard for areas covered in native vegetation (Chambers et al. 2007, BLM 1999 and 2000, Devine 1993). Strategies to reduce the costs of wildfire in the Great Basin involve landscape-level vegetation management programs to prevent and slow the invasion of annual grasses and to restore to desirable vegetation areas that have become dominated by invasive grasses. These vegetation management programs are expensive, and the resources are limited. However, because the spread of invasive grasses is increasing, delays in applying treatment programs will increase future wildfire costs. To assist in decisions regarding how much, when, and where to invest in vegetation management programs, it would help to have information quantifying the benefits of programs that would (1) prevent further spread of invasive grasses to limit the frequency of wildfires to current levels, and (2) restore areas that have been dominated by invasive grasses so that the incidence of wildfire is reduced from current levels. In this paper, we use data from a pilot study for a study to estimate WTP for such programs.

The data used in this paper were collected through a 2005 mail survey of Nevada residents, using survey methods based on Dillman (2000). Focus groups, questionnaire development, and pretesting proceeded through spring and summer of 2005. Responses were analyzed during one-on-one interviews with pretest respondents, during group sessions, and afterward by the researchers. Question wording was reviewed for comprehension and interpretation. The questionnaire collected data necessary to estimate WTP for vegetation management programs to protect ecosystem-derived values. Questions were asked about respondents' demographic characteristics, how they benefit from Great Basin lands, their beliefs and attitudes regarding the effects of invasive annual grasses and wildfire on these lands, and value of land management practices that target invasive grasses and wildfire. Surveys were mailed to 2,125 individuals. The first mailing was conducted during mid-October 2005.

Follow-up postcards were sent out to nonrespondents. A second mailing to nonrespondents was sent during the first week of February 2006, after the holiday season. Of 2,125 surveys sent out, 178 were undeliverable (no forwarding addresses available), and 576 completed surveys were returned for a statewide response rate of 30 percent. Eighteen surveys were eventually omitted from the valuation analysis because they included inconsistent responses (i.e., a 'no' to a bid amount that was lower than one with a 'yes' response from the same respondent).

An experimental design included five survey versions, assigned randomly to participants, to measure the effects of (1) preemptive versus restoration treatment scenarios, (2) providing respondents with additional information, and (3) multiple contingent valuation bid formats.

The contingent valuation question was presented in the context of one of two treatment scenarios. The prevent loss (*PL*) scenario states that the numbers of wildfires in Nevada are expected to double over the next five years due to the continued spread of cheatgrass, increasing the risk of irreversibly losing lands that could support native vegetation. In this case, the proposed vegetation management program would prevent a negative change from the status quo. The obtain gain (*OG*) scenario states that the proposed vegetation management program would be restorative and thereby reduce the number of wildfires throughout Nevada by half. Thus the *OG* scenario measures willingness to pay to improve the status quo. The two versions of the program proposal are provided in Appendix 1. Both scenarios state that the program would be funded through a dedicated tax. Respondents were asked how they would vote in a referendum to implement the vegetation management program where passage would cost each household a specific amount every year for the foreseeable future.

We included in the survey a two-page information sheet about invasive annual weeds (cheatgrass), the increasing costs of wildfire suppression, the accelerating fire cycle, and resulting irreversible ecosystem losses. This information sheet was omitted from a subsample of the multiple-bid *OG* versions in order to investigate the influence of information provision on WTP for fuel management programs.

Three discrete choice CV formats were employed. Respondents were presented with three to nine bid amounts. All versions included the bid amounts \$0 and \$1. The "single-bid" (*S*) version included one additional randomly assigned bid amount from the bid set \$12, \$31, \$52, \$83, \$114, \$157, and \$282. The "double-bid" (*D*) version included two additional randomly assigned bid amounts from the same bid set. The "multiple-bid" (*M*) version included all bid amounts, listed in ascending order. The "no cost" (\$0 bid) and the \$1 bid amounts were included to separate those who are opposed to the proposed program and would vote against it even if it cost them nothing, from those who are in favor of the proposal but are unwilling or unable to pay for it. Following Alberini, Boyle, and Welsh (2003) and Welsh and Poe (1998), all three versions used polychotomous choice response options including 'definitely no,' 'probably no,' 'probably yes,' 'definitely yes,' and 'not sure,' allowing respondents to indicate qualitative levels of response intensity. For reference, an M-version of the discrete choice CV questionnaire is provided in Appendix 2.

Table 1 summarizes variables generated from the survey data that are used in all models. These include variables that indicate 1) respondents' perceptions of the importance of and threat to specific resources and services provided by Nevada's rangelands, 2) what they believe management priorities should be, and 3) attitudes toward management methods. Variable descriptions and measurement units are also listed in Table 1. The first eleven variables are self-explanatory. The following fourteen dummy variables describe the respondent's employment. *Lrt* indicates whether the respondent lives in a large rural town. *Import\_airwat* and *Pri\_grzn* indicate the importance of managing lands to protect air and water quality and for livestock forage, coded on a four-point scale where 1 is not important and 4 is very important. Three dummy variables indicate whether the respondent indicated that rangelands were threatened from overly strict regulations (*Threat\_strict*); cheatgrass (*Threat\_cgrass*); and mining activity (*Threat\_mine*). Four dummy variables indicate whether the respondent disapproved of certain rangeland management methods. These are *Dmthd\_nonna-*

*tive* (seeding with nonnative species to compete with unwanted invasives); *Dmthd\_nogrz* (excluding domestic livestock); *Dmthd\_prsgrz* (use of livestock for prescribed grazing to reduce cheatgrass accumulation); and *Dmthd\_handcut* (labor intensive hand cutting and removal of brush to reduce fuel loads). Two variables indicate respondents' general attitudes toward rangeland wildfire control: *Dfire\_all* indicates whether the respondent felt strongly that all rangelands fires should be suppressed, and *Dfire\_humanlife* indicates the respondent felt strongly that rangeland fires should be allowed to burn unless they threaten human life.

### Estimation Models and Response Recoding Procedure

#### Multinomial Logit

Individual predicted response probabilities (DY, PY, PN, DN, and NS) to each bid amount are estimated using a multinomial logit, with NS as the base outcome. The specification is general enough to account for several reasons for expressed response intensity.<sup>10</sup> Let  $\text{Prob}(\text{RESPONSE}_{ijk})$  denote the probability that individual *i* presented with bid amount *j* chooses response *k* where  $k = \text{DY, PY, PN, DN, and NS}$ . The multinomial logit model includes the following four groups of explanatory variables: a bid amount presented ( $b_j$ ); a vector of respondent demographic variables ( $x_i$ ); a vector of variables characterizing respondents' perceptions as described in the previous section ( $z_i$ ); and a vector of variables that characterize the survey

<sup>10</sup> An ordered logit might appear to be a logical specification in estimating response probabilities, since distributions for WTP and bid amount could be expected to move in one direction only. We explored this possibility and found that estimation results do not seem to be ordered purely by "Bid," but instead it appears that NS may not fall between PY and PN. We estimated model (1) using an ordered logit with the ordering moving along the continuum from DN through DY with NS in the middle and with the order of NS and PN switched. There was almost no difference in goodness of fit between the two models, so we cannot rule out that the NS and PN categories are not different. However, compared with the multinomial logit results, the covariates influence these two categories quite differently. While the ordering appears to work well for the other categories, the results on the NS category are not consistent with a simple ordering. Moreover, exploratory research suggests that the connection of this category to the covariates is not what one would expect from a simple ordering. It is likely that this is a mixed category that needs more detailed exploration in future research. For these reasons, we proceed with the multinomial logit specification, where the structural restrictions on the data are minimal.

**Table 1. Variables Describing Bid Responses, Demographics, and Survey Versions**

Variable	Definition	Mean	St. Dev.
<i>Bid</i>	Dollar amount presented to respondent	71.974	85.089
<i>Vs</i>	1 = single-bid questionnaire version; 0 otherwise	0.103	0.305
<i>Vd</i>	1 = double-bid questionnaire version; 0 otherwise	0.099	0.298
<i>Vm</i>	1 = multi-bid questionnaire version; 0 otherwise	0.798	0.402
<i>OG</i>	Scenario: 1 = restoration (obtain gain); 0 = prevent loss	0.631	0.483
<i>Info</i>	1 = information sheet provided; 0 = no information sheet	0.626	0.484
<i>Income</i>	Household annual income in \$1000s	71.539	53.189
<i>Hshld</i>	Number of people in household	2.457	1.251
<i>Age</i>	Age of respondent	52.647	14.752
<i>Educ</i>	Number of years of completed schooling	14.477	2.421
<i>Yrs_NV</i>	Number of years lived in Nevada	21.705	12.216
<i>Job_ag</i>	1 = ranching or agriculture; else = 0	0.074	0.262
<i>Job_Indscp</i>	1 = landscaping; else = 0	0.019	0.135
<i>Job_mine</i>	1 = mining; else = 0	0.144	0.352
<i>Job_trade</i>	1 = wholesale or retail trade; else = 0	0.081	0.274
<i>Job_wtrmgnt</i>	1 = water resources management; else = 0	0.026	0.159
<i>Job_othutil</i>	1 = utilities (other than water); else = 0	0.022	0.148
<i>Job_health</i>	1 = healthcare; else = 0	0.081	0.274
<i>Job_nrsci</i>	1 = natural resource/environmental sciences; else = 0	0.052	0.222
<i>Job_ed</i>	1 = education/academia; else = 0	0.100	0.301
<i>Job_ent</i>	1 = arts, entertainment, hotel, food services; else = 0	0.056	0.229
<i>Job_recr</i>	1 = outdoor recreation & tourism; else = 0	0.063	0.243
<i>Job_publnds</i>	1 = public land management; else = 0	0.044	0.206
<i>Job_admin</i>	1 = public admin (not land or water resources); else = 0	0.015	0.121
<i>Job_fire</i>	1 = firefighting; else = 0	0.015	0.121
<i>Lrt</i>	1 = lives in large rural town; else = 0	0.130	0.337
<i>Import_airwat</i>	Importance of air/water quality (1-4)	3.642	0.627
<i>Threat_strict</i>	1 = strict regulations threaten rangelands; else = 0	0.464	0.500
<i>Threat_cgrass</i>	1 = cheatgrass spread threatens rangelands; else = 0	0.849	0.359
<i>Threat_mine</i>	1 = mining threatens rangelands; else = 0	0.434	0.497
<i>Pri_grzn</i>	Importance of livestock as a management priority (1-4)	2.232	1.003
<i>Dmthd_nonnative</i>	1 = seeding nonnative species <i>not</i> appropriate; else = 0	0.276	0.448
<i>Dmthd_nogrz</i>	1 = excluding grazing animals <i>not</i> appropriate; else = 0	0.373	0.485
<i>Dmthd_prsgrz</i>	1 = prescribed grazing <i>not</i> appropriate; else = 0	0.052	0.223
<i>Dmthd_handcut</i>	1 = brush/tree cutting by hand <i>not</i> appropriate; else = 0	0.049	0.215
<i>Dfire_all</i>	1 = all rangeland fires should be stopped; else = 0	0.524	0.500
<i>Dfire_humanlife</i>	1 = only stop rangeland fires that threaten human life	0.457	0.499



version assigned to individual  $i$  ( $d_i$ ). For more details about the explanatory variables, see the Results section.

In addition to a \$0 bid and a \$1 bid, individuals were presented with one, two, or seven bid amounts depending on the questionnaire version they received (the  $S$ ,  $D$ , or  $M$  questionnaire versions). Therefore, the specification results in unbalanced panel estimation.<sup>11</sup> In estimating this model and all subsequent models, a subsample of the data is used in which respondents identified as "opposers" and "protesters" are dropped. Respondents who voted 'definitely no' to support the program at zero cost ("opposers") are dropped, while those who did not vote DN if it cost them nothing ("non-opposers") were retained. The omission of this part of the sample means that the WTP estimates apply only to that part of the population for whom the proposal is not welfare reducing. A battery of questions asking why respondents voted the way that they did to valuation proposals was included in the questionnaire to identify protest responses. Respondents were coded as "protesters" if they checked any of the following four reasons for voting 'no' at least once to any dollar bid amount: "I don't trust the government to use my taxes wisely," "I already pay too much in taxes," "I object to the way the question was asked," or "I feel that I don't have enough information." The discussions of treatment of opposition and protest responses for this data set are found elsewhere (Rollins et al. 2009). The resulting subsample contains 1,577 observations from 233 respondents.

#### *Alternative Definitions of 'Yes' Responses*

The next question is how to map five response categories onto a bivariate 'yes' or 'no' response variable, to most effectively use the information contained in the data to estimate WTP. For comparison's sake, we start with two approaches represented in the literature. One method is to code actual DY responses (or alternatively, all DY

+ PY responses) as 'yes' and all others as 'no,' effectively taking all responses at their face value. The rules are applied to all of the response categories and do not account for individual differences in how the intensity of 'yes' responses may be expressed.

A second method is an approach similar to that of Caudill and Groothuis (2005) (hereafter we denote this approach as C&G). Using data with three response options ('yes,' 'no,' and DK) and a variety of covariates, they used a multinomial logit specification to estimate the individual probability of each response option. They recoded each DK response to either 'yes' or 'no' depending on which of the three response categories had the highest predicted probability. We adapt this approach for five response categories by recoding each PY, PN, and NS response as 'yes' if the predicted response probability for DY,  $p^{\wedge}(DY)$ , is higher than for any other response category, and all others are coded as 'no.'

These two examples are representative of existing methods of treating data in that they rely on some mechanism for recoding actual responses that indicate a 'yes' direction. This is in contrast to the alternative method we present here, where we include into the 'yes' category additional responses from the categories where weaker response intensity is expressed (i.e. PY, PN, and NS) according to the multinomial logit results. The rule we take is that responses with predicted probabilities of choosing DY above a series of predicted DY cutoffs form a series of data sets (for example, one data set could be defined where 'yes' is all actual DY plus observations from other categories with a predicted probability of choosing DY of over 95 percent; the next where the cutoff is predicted DY probability of over 90 percent; and so on). In order to demonstrate how WTP changes with successive definitions of the cutoff, we use 10 percent increments to create a series of data sets with which to estimate a series of WTP point estimates.

#### *WTP Estimation*

We then estimate WTP for each individual data set, as defined by the level of predicted DY response probability. In dealing with an unbalanced panel in which each respondent is

<sup>11</sup> Individual effects from the multiple responses per individual are accounted for by relaxing the usual requirement that the observations be independent. That is, observations are independent across individuals (clusters) but not necessarily within individuals. This is accommodated in Stata by using the "vce option" (Stata Data Analysis and Statistical Software, StataCorp LP, College Station, TX).

represented by three to nine observations with one observation per bid amount, we use random effects probit as described in Rollins, Dumitras, and Castledine (2008) and Boxall, Rollins, and Englin (2003). Let  $\text{Prob}(Yes_{ij})$  denote the probability that individual  $i$  presented with bid amount  $j$  chooses a 'yes' response. The random effects probit model includes four sets of explanatory variables as in the multinomial logit model. While the nature of the explanatory variables is the same as in the multinomial logit model, the specific variables included in the WTP estimation are different. In the random effects probit model,  $b_j$  is *Bid*;  $x_j$  includes *Income*, *Age*, *Age*<sup>2</sup>, *Yrs\_NV*, *Job\_ag*, *Job\_indscp*, *Job\_mine*, *Job\_constr/mfn*, *Job\_trade*, *Job\_wtrmgnt*, *Job\_othutil*, *Job\_health*, *Job\_nrscl*, *Job\_ed*, *Job\_ent*, *Job\_recr*, *Job\_publnds*, *Job\_admin*, *Job\_fire*, and *Lrt*;  $z_j$  includes *Dmthd\_prsgrz*; and  $d_j$  includes *Vs*, *Vd*, *OG*, and *Info*. By dropping "opposer" observations, we measure conditional WTP ( $WTP|WTP \geq 0$ ).

## Results

### *Multinomial Logit*

The multinomial logit model is estimated with a total of 1,440 observations from 220 respondents, resulting in a pseudo R-squared of 0.3102. Table 2 lists the specific explanatory variables included in the model and summarizes the estimation results. For ease of interpretation, only marginal effects are presented.

As expected, *ceteris paribus*, a higher bid amount is associated with a higher probability of responding DN or PN, while it is associated with a lower probability of DY responses. However, the marginal effect of a bid amount (*Bid*) on the probability of NS responses is also positive, and the magnitude is as large as that for DN responses. The marginal effect on PY is statistically not different from zero. This result also suggests that responses are not ordered as one would expect (as DN, PN, NS, PY, and then DY) with NS representing the middle of the response intensity spectrum.

*Income* is positively associated with the probability of DY responses. By contrast, *Income* has a negative effect on the probability of choosing the NS and DN response categories. Household size

(*Hshld*) and number of years living in Nevada (*Yrs\_NV*) have a positive and negative influence on the likelihood of choosing DY, respectively. In terms of the effect of *Yrs\_NV*, people who have lived in the area for a longer time have experienced wildfires and the gradual encroachment of annual grasses that fuel them. Consequently, they may be somewhat inured to frequent wildfires and have ambivalent feelings as to the urgency of taking costly steps to address the problem, relative to newcomers to the state. Education (*Educ*) tends to decrease the probability of choosing PN and PY, while it is positively associated with NS probability. This effect is consistent with the interpretation that more educated people may have a greater tendency to frame and consider public policy issues in a more complex manner, in which they see a need to weigh a variety of pros and cons (Miller, Slomczynski, and Kohn 1985, Kohn and Schooler 1978).

Version effects include 1) the management program version effect, where the sample was split between obtain gain (*OG*) and prevent loss (*PL*) versions; 2) the information version effect (*Info*); and 3) bid number version effects (*Vs* and *Vd*). Those receiving the *OG* version are more likely to vote DN relative to those with the *PL* version of the referendum question, implying that people are likely to be willing to pay more for a prevention program than a program to restore ecosystem losses from invasive weeds and wildfires in the Great Basin. The sign on *OG* for NS responses is negative (but not significant), implying that this version effect also dampens the probability of a NS response. The subsample of respondents with additional information (*Info*) had a lower probability of choosing NS, implying that the information influenced respondents to choose one of the other four response categories (all of which indicate stronger intensity of attitude) when they otherwise would have answered NS. That is, more information decreases the probability of expressing 'not sure.' Relative to those assigned with multiple dollar bid amounts, respondents with the double bid amount versions (*Vd*; \$0, \$1, plus two other bids) were not different in response patterns. However, those who received the single bid versions (*Vs*; \$0, \$1, plus one other bid) were less likely to choose DN and NS.

**Table 2. Marginal Effects on Response Probabilities of Explanatory Variables** (*Multinomial Logit*)

	(1) DN	(2) PN	(3) NS	(4) PY	(5) DY
<i>Bid</i>	0.002*** (0.000)	0.001** (0.000)	0.002*** (0.000)	0.001 (0.000)	-0.005*** (0.001)
<i>Income</i>	-0.001*** (0.000)	0.000 (0.000)	-0.001* (0.001)	0.000 (0.000)	0.002** (0.001)
<i>Educ</i>	0.000 (0.009)	-0.019* (0.010)	0.026* (0.015)	-0.026* (0.014)	0.018 (0.017)
<i>Hshld</i>	-0.020 (0.017)	0.002 (0.015)	-0.035 (0.025)	0.003 (0.019)	0.050* (0.027)
<i>Yrs_NV</i>	0.002 (0.002)	-0.001 (0.001)	0.002 (0.002)	0.001 (0.002)	-0.004* (0.003)
<i>Job_ag</i>	-0.015 (0.058)	0.031 (0.107)	-0.069 (0.058)	0.245** (0.117)	-0.191** (0.092)
<i>Job_mine</i>	-0.081** (0.033)	0.056 (0.071)	0.055 (0.109)	-0.115* (0.067)	0.086 (0.102)
<i>Job_trade</i>	-0.106*** (0.026)	-0.042* (0.024)	-0.090 (0.071)	0.273** (0.108)	-0.034 (0.125)
<i>Job_nrsci</i>	0.392 (0.328)	-0.038 (0.032)	-0.149*** (0.045)	0.065 (0.240)	-0.271** (0.107)
<i>Job_ed</i>	-0.072** (0.031)	0.136 (0.155)	0.073 (0.134)	-0.169*** (0.060)	0.032 (0.140)
<i>Job_recr</i>	0.010 (0.080)	-0.001 (0.058)	0.051 (0.170)	-0.159*** (0.056)	0.099 (0.174)
<i>Job_publnds</i>	-0.086** (0.038)	-0.030 (0.022)	0.137 (0.223)	-0.051 (0.097)	0.030 (0.179)
<i>Lrt</i>	0.203** (0.083)	-0.027 (0.026)	-0.171*** (0.037)	0.041 (0.073)	-0.046 (0.087)
<i>Import_airwat</i>	0.050 (0.037)	-0.015 (0.020)	-0.078** (0.038)	-0.036 (0.048)	0.079 (0.055)
<i>Threat_strict</i>	-0.071* (0.041)	-0.059 (0.047)	0.189*** (0.065)	-0.122** (0.048)	0.063 (0.068)
<i>Threat_cgrass</i>	0.063* (0.038)	-0.072 (0.065)	-0.102 (0.095)	0.065 (0.064)	0.046 (0.100)
<i>Threat_mine</i>	-0.018 (0.037)	-0.091** (0.041)	-0.038 (0.051)	-0.037 (0.049)	0.185*** (0.063)
<i>Pri_grzn</i>	-0.011 (0.017)	-0.007 (0.015)	-0.012 (0.029)	-0.052* (0.030)	0.083** (0.037)
<i>Dmthd_nonnative</i>	0.015 (0.039)	-0.033 (0.028)	0.099 (0.069)	-0.084 (0.056)	0.002 (0.073)
<i>Dmthd_nogrzn</i>	-0.015 (0.033)	-0.021 (0.028)	0.091 (0.059)	-0.074 (0.050)	0.019 (0.068)
<i>Dmthd_prsgrzn</i>	0.045 (0.117)	0.106 (0.125)	0.153 (0.161)	-0.174** (0.072)	-0.130 (0.166)
<i>Dmthd_hndcut</i>	0.324** (0.154)	0.125 (0.133)	-0.160*** (0.035)	-0.186*** (0.057)	-0.102 (0.132)
<i>Dfire_all</i>	0.116*** (0.039)	0.037 (0.025)	-0.114* (0.060)	-0.020 (0.049)	-0.019 (0.065)
<i>Dfire_humanlife</i>	0.095** (0.040)	0.012 (0.042)	-0.057 (0.052)	0.064 (0.047)	-0.114* (0.067)
<i>Vs</i>	-0.081* (0.042)	-0.034 (0.028)	-0.096* (0.058)	0.133 (0.097)	0.079 (0.117)
<i>Vd</i>	0.057 (0.103)	-0.039 (0.029)	-0.089 (0.074)	-0.032 (0.089)	0.104 (0.127)
<i>OG</i>	0.108** (0.045)	0.011 (0.033)	-0.085 (0.100)	0.048 (0.075)	-0.083 (0.129)
<i>Info</i>	0.035 (0.038)	0.067 (0.043)	-0.133* (0.076)	0.041 (0.065)	-0.009 (0.090)

Notes: Robust standard errors in parentheses. Significance levels of 0.01, 0.05, and 0.1 denoted by three, two and one asterisks (\*\*\*, \*\*, \*) respectively.

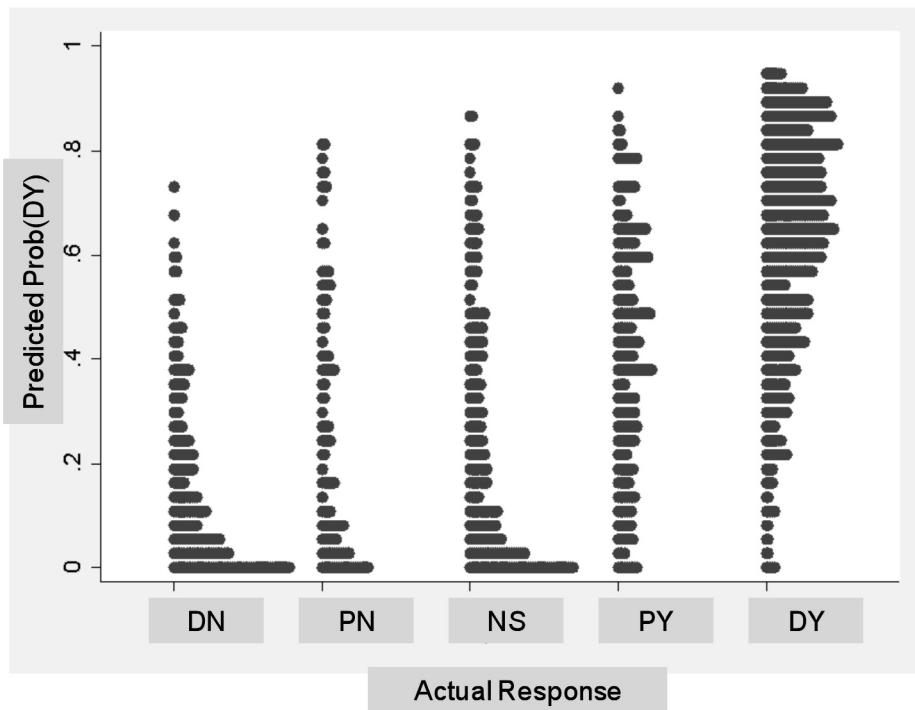
*Distributions of Predicted Response Probabilities by Actual Response Categories*

Using the multinomial logit parameter estimates, predicted probabilities for each response category are calculated for each observation. Figure 1 illustrates the predicted probabilities of DY responses for each of the five actual response categories. We see a mass of observations in the lower portion (lower predicted probabilities for DY responses) for the actual DN responses, and another mass on the upper portion (higher predicted probabilities for DY responses) for the actual DY responses. This indicates that when high intensity is expressed (i.e., ‘definitely’) for either response direction (‘yes’ or ‘no’), predicted direction and intensity in terms of predicted probabilities of DY and DN agree with the actual responses for most observations. Thus, predicted and actual DY and DN responses are so closely associated that using the predicted probabilities would add little information. Accordingly, we treat actual DY and DN responses as their face value in subsequent

analyses. The middle three columns of Figure 1 show, for all actual PN, NS, and PY responses, the predicted probability of being a DY instead, based on all covariates.

*Effect of Alternative Definitions of ‘Yes’ Responses on WTP Estimates*

Following Hanemann (1984), WTP is estimated for each individual using data sets generated by the three approaches described above. The first row of Table 3 represents WTP estimates generated from defining ‘yes’ as actual DY responses only (all other responses are coded as ‘no’). The second row presents WTP when ‘yes’ is defined as both actual DY and PY responses. The corresponding estimates of mean WTP are \$39.83 and \$94.22. The third row results from our adaptation of Caudill and Groothuis (2005) (C&G), with an estimated mean WTP of \$80.71. The leftmost points on Figure 2 illustrate these first three estimates of mean WTP and 95 percent confidence intervals. The remaining nine rows of Table 3



**Figure 1. Distribution of Predicted Response Probabilities by Actual Response**

show the series of WTP estimates under the alternative definitions of 'yes' that we propose in this paper.<sup>12</sup>

The C&G approach generated a relatively high mean WTP estimate (\$80.71) because it codes as 'yes' every low-intensity observation (PY, PN, NS) if its highest predicted probability is for DY, regardless of the magnitude of predicted probability of DY. Given five response categories, this means that an observation with a predicted probability of DY as low as 21 percent could be coded as 'yes' (where the probabilities of the other response categories are each less than 20 percent), as long as DY is the category with the highest predicted probability of being selected by the individual.<sup>13</sup> In contrast, our approach controls for the level of predicted probability of a DY response by setting allowable cutoffs for the predicted DY probability as a rule for recoding

a response as a 'yes' response. While C&G compare across all categories in identifying the one with the highest probability (and then code as 'yes' if the highest category is DY), our approach does not compare across categories, but only uses the predicted probability of DY. Note that the information about other response categories is not discarded; it is used in the estimation of predicted probabilities, and the values of  $p^{\wedge}(DY)$  we use in response recoding reflect information about all the other response categories.

In contrast to "one size fits all" block recoding of low-intensity responses as under previous methods, our approach allows assessment of the sensitivity of WTP estimates to specific and known levels of the probability of a DY response. By using different cutoff levels to code responses to 'yes' or 'no,' we can observe the resulting range of mean WTP estimates. This provides very different information about ranges of WTP than is supplied by confidence intervals around mean WTP estimates generated with a single data set, as is the case in other methods. In our case, as seen in Figure 2, mean WTP estimates appear relatively stable up to the 'yes' definition of  $DY + p^{\wedge}(DY) \geq 70$  percent. At the 70 percent level,

<sup>12</sup> When 'yes' is defined as  $DY + p^{\wedge}(DY) \geq 0\%$ , so that all PY, PN, and NS responses are coded as 'yes,' the mean WTP is \$289.06. For scaling purposes, we excluded this level from Figure 2.

<sup>13</sup> In C&G's original paper, they used three categories, not five, so that the lowest predicted probability of DY could have been 34 percent. On the other hand, had four categories been used, the lowest predicted probability could have been 26 percent.

**Table 3. WTP Estimates with Alternative Definitions of 'Yes' Responses**

	Definition of 'yes'	Obs.	%Yes	Resp.	Mean	Median	95% CI
(1)	DY only	1577	39%	233	39.83	39.35	(35.58, 44.09)
(2)	DY + PY	1577	55%	233	94.22	89.44	(88.55, 99.89)
(3)	C&G	1577	53%	233	80.71	78.56	(75.76, 85.65)
(4)	$DY + p^{\wedge}(DY) \geq 80\%$	1577	40%	233	42.96	39.94	(38.49, 47.43)
(5)	$DY + p^{\wedge}(DY) \geq 70\%$	1577	41%	233	47.58	44.04	(43.1, 52.06)
(6)	$DY + p^{\wedge}(DY) \geq 60\%$	1577	44%	233	56.97	52.91	(52.44, 61.51)
(7)	$DY + p^{\wedge}(DY) \geq 50\%$	1577	48%	233	66.04	62.59	(61.15, 70.92)
(8)	$DY + p^{\wedge}(DY) \geq 40\%$	1577	52%	233	79.38	79.32	(74.32, 84.44)
(9)	$DY + p^{\wedge}(DY) \geq 30\%$	1577	56%	233	94.14	88.82	(88.39, 99.88)
(10)	$DY + p^{\wedge}(DY) \geq 20\%$	1577	59%	233	103.86	100.97	(98.34, 109.37)
(11)	$DY + p^{\wedge}(DY) \geq 10\%$	1577	64%	233	116.10	112.32	(110.81, 121.39)
(12)	$DY + p^{\wedge}(DY) \geq 0\%$	1577	75%	233	289.06	269.69	(269.18, 308.94)

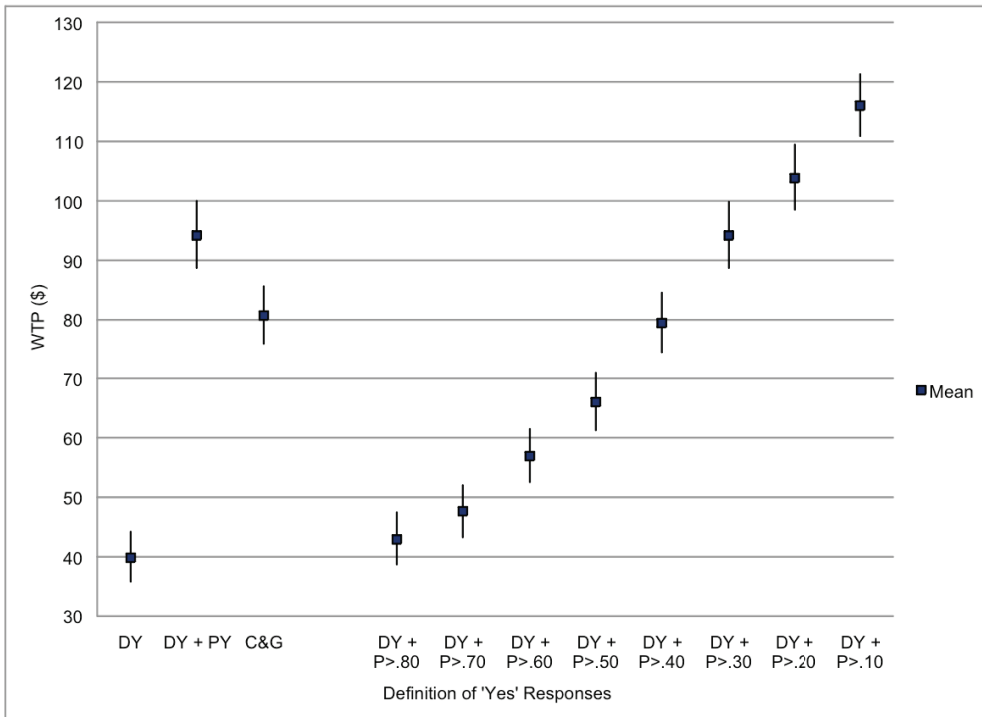
Note: 'Definitely no' responses are always treated as 'no.'

mean WTP is \$47.58, which is 20 percent higher than the most conservative estimate of \$39.83 (DY only). The 95 percent confidence intervals around WTP estimates for the ‘yes’ definitions of DY only,  $DY + p^{\wedge}(DY) \geq 80$  percent, and  $DY + p^{\wedge}(DY) \geq 70$  percent are all overlapping, while the move from  $DY + p^{\wedge}(DY) \geq 70$  percent to  $DY + p^{\wedge}(DY) \geq 60$  percent entails a change in the shape (steeper) of the plotted series of mean WTP estimates, and no overlapping of confidence intervals with the next lower WTP estimate. We can contrast these results with the results under the C&G method that produces a single mean WTP estimate. The C&G mean WTP (\$80.71) is greater than the mean WTP under the ‘yes’ definition of  $DY + p^{\wedge}(DY) \geq 40$  percent (\$79.38). Our method makes an important contribution by generating a series of WTP estimates over specified levels of probability that respondents choose ‘yes’ with the highest intensity. This allows a more complete interpretation of the range of WTP estimates.

**Conclusions and Further Research**

As Caudill and Groothuis (2005) point out, much of the literature responds to the problem of mapping multiple response categories into a binary variable as an empirical issue. Our approach also treats this as an inductive empirical problem but is less restrictive than others, allowing the data to provide more information. We demonstrate that in attempting to account for random measurement errors commonly encountered in measuring response intensity, and to account for errors due to response mapping rules chosen by the researcher, a range of WTP estimates can be produced that spans a continuum of lower bounds placed on the predicted probability of a DY (‘definitely yes’) response. The shape of the plotted series of resulting WTP estimates reflects regions in which WTP is more and less sensitive to changes to alternative lower limits on the predicted probability of a DY.<sup>14</sup>

<sup>14</sup> Confidence intervals on each WTP point estimate are artifacts of underlying assumptions regarding the degree to which one is comfortable taking each response at its face value.



**Figure 2. Estimated WTP under Alternative Definitions of ‘Yes’ Responses: Mean and 95% Confidence Interval**

One advantage of the approach taken in this paper is that it leads us away from block recoding/interpreting of specific response choices to understanding that true DY respondents may choose *any* of the response choices with a non-zero probability. Estimating DY probabilities using both the CV measures and a substantial array of control variables that could condition response tendencies and other aspects of response intensity unrelated to preference uncertainty enables us much more closely to approximate the true underlying attitude of each respondent. This in turn helps clarify that the mapping of these probabilistic responses into a dichotomous variable ('yes' vs. 'no') necessarily involves choosing a cutoff or displaying answers as ranges. At this point, with as much as we know, the cutoff choice remains a matter of judgment, but future research will be able to compare and contrast the linkages of different cutoff choices or ranges to a variety of measured causes and consequences. This inductive procedure will ultimately lead us towards a new customary set of cutoff points or ranges in this area of research.

The multinomial logit results for the response categories of NS ('not sure') and PN ('probably no') categories suggest that there may be a mixture of distributions in the NS category. The survey question was formatted such that the NS category was set off to one side on the questionnaire, and not as a middle option. The NS responses may have captured some genuine middle responses [as suggested by Wang (1997), where the presented bid is very close to the actual WTP for the respondent] as well as others that would better be considered a separate response category altogether. An alternative format might include a middle option as well as a separate category to capture the non-central NSs. Separating these response categories might have resulted in better predicted values for the probability of DY attitudes.

The method presented in this paper produces a range of WTP estimates, with each estimate conditioned on the cutoff level for predicted probability of a DY response. While decreasing cutoff levels would generally be associated with higher WTP point estimates, the shape of the plotted series of WTP estimates for a given data set is

purely an empirical issue. In our case, treating observations with a predicted probability of DY of up to 70 percent as 'yes' led to WTP estimates with confidence intervals that still overlapped those from using DY-only observations as 'yes.' Beyond this, the WTP estimates diverged considerably. With data from other applications, the ranges over which WTP estimates exhibit overlapping confidence intervals might be wider, and the plot of WTP point estimates much flatter.

This approach could be improved upon to even more effectively use information contained in the data. For example, by using successive cutoff levels of predicted probability of a DY response, our method produces a range of WTP estimates. However, because the series of WTP point estimates results from different data sets, it is not a distribution of WTP in a statistical sense. Research to generate a statistical distribution for WTP estimates, with the associated statistical properties of that distribution, would be a useful extension for future work.

The take-home message for policy-making is that the choice of criteria for coding response intensity for observations that are used to derive WTP estimates can have dramatic implications on estimated WTP. In the case of WTP for invasive annual grass treatments and restoration used in the application for this paper, decision makers may wish to err on the side of caution with regard to treatment, because the opportunity cost of delaying treatment entails an even higher rate of spread of invasive grasses and larger future firefighting costs. A more complete picture of how WTP can vary with changes in the criteria for treating intensity of a 'yes' response may be desirable in situations where the opportunity costs of making an error with a more stringent rule (i.e., stated DY responses only coded as 'yes') are much greater than those of a less stringent rule (e.g., a cutoff at 70 percent predicted probability of a DY to be coded as 'yes'). One practical application of this method would be to use the ranges of WTP estimates in sensitivity analyses for benefit cost analysis. The alternative of using a single point estimate and confidence intervals based on any of the other methods would seem to be underutilizing the available information contained in the data.

## References

- Aadland, D., and A.J. Caplan. 2006. "Curbside Recycling: Waste Resource or Waste of Resources?" *Journal of Policy Analysis and Management* 25(4): 855–874.
- Akter, S., R. Brower, L. Brander, and P. Van Beukering. 2009. "Respondent Uncertainty in a Contingent Market for Carbon Offsets." *Ecological Economics* 68(6): 1858–1863.
- Albaum, G., C. Roster, J.H. Yu, and R.D. Rogers. 2007. "Simple Rating Scale Formats: Exploring Extreme Response." *International Journal of Market Research* 49(5): 1–17.
- Alberini, A., K. Boyle, and M. Welsh. 2003. "Analysis of Contingent Valuation Data with Multiple Response Options Allowing Respondents to Express Uncertainty." *Journal of Environmental Economics and Management* 45(1): 40–62.
- Arieli, D., G. Loewenstein, and D. Prelec. 2006. "Tom Sawyer and the Construction of Value." *Journal of Economic Behavior and Organization* 60(1): 1–10.
- Bateman, I.J., D. Burgess, and W.G. Hutchinson. 2008. "Learning Design Contingent Valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness." *Journal of Environmental Economics and Management* 55(2): 127–141.
- Bishop, G.F. 1987. "Experiments with the Middle Response Alternative in Survey Questions." *The Public Opinion Quarterly* 51(2): 220–232.
- Blumenschein, K., G. Blomquist, M. Johannesson, N. Horn, and P. Freeman. 2008. "Eliciting Willingness to Pay without Bias: Evidence from a Field Experiment." *The Economic Journal* 118(525): 114–137.
- Bollen, K.A. 1989. *Structural Equations with Latent Variables*. New York: Wiley-Interscience.
- Boxall, P., K. Rollins, and J. Englin. 2003. "Heterogeneous Preferences for Congestion during a Wilderness Experience." *Resource and Energy Economics* 25(2): 177–195.
- Broberg, T., and R. Brannlund. 2008. "An Alternative Interpretation of Multiple Bounded WTP Data—Certainty Dependent Payment Card Intervals." *Resource and Energy Economics* 30(4): 555–567.
- Bureau of Land Management (BLM). 1999. "Out of Ashes, an Opportunity." National Interagency Fire Center, Boise, ID, p. 36. Available at <http://www.blm.gov/pgdata/etc/medialib/blm/id/fire/gbri/documents.Par.5395.File.dat/Ashes.pdf> (accessed January 11, 2010).
- Bureau of Land Management (BLM). 2000. "The Great Basin: Healing the Land." National Interagency Fire Center, Boise, ID, p. 28. Available from Bureau of Land Management, 1340 Financial Blvd, Reno, Nev, 89520-0006 (775/861-6400).
- Cameron, T., and J. Englin. 1997. "Welfare Effects of Changes in Environmental Quality under Individual Uncertainty about Use." *Rand Journal of Economics* 28(0): S45–70.
- Cantril, H. 1946. "The Intensity of an Attitude." *Journal of Abnormal and Social Psychology* 41(2): 129–135.
- Carroll, R., J. Carroll, D. Ruppert, and L.A. Stefanski. 2006. *Measurement Error in Nonlinear Models: A Modern Perspective* (2nd edition). New York: Chapman & Hall/CRC.
- Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell, S. Presser, P.A. Ruud, V.K. Smith, M. Conaway, and K. Martin. 1998. "Referendum Design and Contingent Valuation: The NOAA Panel's No-Vote Recommendation." *The Review of Economics and Statistics* 80(2): 335–338.
- Chambers, J., B.A. Roundy, R.R. Blank, S.E. Meyer, and A. Whitaker. 2007. "What Makes Great Basin Sagebrush Ecosystems Invasible by *Bromus tectorum*?" *Ecological Monographs* 77(1): 117–145.
- Champ, P.A., and R.C. Bishop. 2001. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias." *Environmental and Resource Economics* 19(4): 383–402.
- Champ, P.A., R.C. Bishop, T.C. Brown, and D.W. McCollum. 1997. "Using Donation Mechanisms to Value Nonuse Benefits from Public Goods." *Journal of Environmental Economics and Management* 33(2): 151–162.
- Champ, P., A. Alberini, and C. Correas. 2005. "Using Contingent Valuation to Value a Noxious Weeds Control Program: The Effects of Including an Unsure Response Category." *Ecological Economics* 55(1): 47–60.
- Crandall, J.E. 1965. "Some Relationships among Sex, Anxiety, and Conservatism of Judgment." *Journal of Personality* 33(1): 99–107.
- Caudill, S., and P. Groothuis. 2005. "Modeling Hidden Alternatives in Random Utility Models: An Application to 'Don't Know' Responses in Contingent Valuation." *Land Economics* 81(3): 445–454.
- DeShazo, J.R. 2002. "Designing Transactions without Framing Effects in Iterative Question Formats." *Journal of Environmental Economics and Management* 43(3): 360–385.
- DeShazo, J.R., and G. Fermo. 2002. "Designing Choice Sets of Stated Preference Methods: The Effects of Complexity on Choice Consistency." *Journal of Environmental Economics and Management* 44(1): 123–143.
- Devine, R. 1993. "The Cheatgrass Problem." *The Atlantic* 271(5): 40–48.
- Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method* (2nd edition). New York: Wiley.
- Elliot, M.N., A.M. Haviland, D.E. Kanouse, K. Hambarsoomian, and R.D. Hays. 2009. "Adjusting for Subgroup Differences in Extreme Response Tendency in Ratings of Health Care: Impact on Disparity Estimates." *Health Research and Educational Trust* 44(2): 542–561.
- Fuller, W.A. 1986. *Measurement Error Models*. New York: Wiley.
- Greenleaf, E.A. 1992. "Extreme Response Style." *The Public Opinion Quarterly* 56(3): 328–351.
- Groothuis, P., and J. Whitehead. 2002. "Does Don't Know Mean No? Analysis of Don't Know Responses in Dichotomous Choice Contingent Valuation Questions." *Applied Economics* 34(15): 1935–1940.
- Groves, R.M. 1991. "Measurement Error across Disciplines." In P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman, eds., *Measurement Errors in Surveys*. New York: Wiley.
- Haener, M., and W. Adamowicz. 1998. "Analysis of 'Don't Know' Responses to Referendum Contingent Valuation Questions." *Agricultural and Resource Economics Review* 27(2): 218–230.



- Hamilton, D.L. 1968. "Personality Attributes Associated with Extreme Response Style." *Psychological Bulletin* 69(3): 192-203.
- Hanemann, W.M. 1984. "Welfare Estimation in Contingent Valuation Experiments with Discrete Responses." *American Journal of Agricultural Economics* 66(3): 332-341.
- Hui, C.H., and H.C. Triandis. 1985. "The Instability of Response Sets." *Public Opinion Quarterly* 49(2): 253-260.
- Kelley, J. (1973). "Causal Chain Models for the Socio-economic Career." *American Sociological Review* 38(4): 481-493.
- Kohn, M.L., and C. Schooler. 1978. "The Reciprocal Effects of the Substantive Complexity of Work and Intellectual Flexibility: A Longitudinal Assessment." *American Journal of Sociology* 84(1): 24-52.
- Krosnick, J.A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5(3): 213-236.
- Krosnick, J.A. 1999. "Survey Research." *Annual Review of Psychology* 50(1): 537-567.
- Krosnick, J.A., C.M. Judd, and B. Wittenbrink. 2005. "The Measurement of Attitudes." In D. Albarraci, B.T. Johnson, and M.P. Zanna, eds., *The Handbook of Attitudes*. Mahwah, NJ: Erlbaum.
- Lewis, N.A., and J.A. Taylor. 1955. "Anxiety and Extreme Response Preferences." *Educational and Psychological Measurement* 15(2): 111-116.
- Li, C., and L. Mattsson. 1995. "Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation." *Journal of Environmental Economics and Management* 28(2): 256-269.
- Loomis, J., and E. Ekstrand. 1998. "Alternative Approaches for Incorporating Respondent Uncertainty when Estimating Willingness to Pay: The Case of the Mexican Spotted Owl." *Ecological Economics* 21(1): 29-41.
- Miller, J., K.M. Slomczynski, and M.L. Kohn. 1985. "Continuity of Learning-Generalization." *American Journal of Sociology* 91(3): 593-615.
- Mozumder, P., and R.P. Berrens. 2007. "Investigating Hypothetical Bias: Induced-Value Tests of the Referendum Voting Mechanisms with Uncertainty." *Applied Economics Letters* 14(10): 705-709.
- Naemi, B.D., D.J. Beal, and S.C. Payne. 2009. "Personality Predictors of Extreme Response Style." *Journal of Personality* 77(1): 261-286.
- NOAA. 1993. "National Oceanographic and Atmospheric Administration, Appendix I: Report of the NOAA Panel on Contingent Valuation." *Federal Register* 58: 4602-4614.
- Ready, R., S. Navrud, and W.R. Dubourg. 2001. "How Do Respondents with Uncertain Willingness to Pay Answer Contingent Valuation Questions?" *Land Economics* 77(3): 315-326.
- Ready, R., J. Whitehead, and G. Blomquist. 1995. "Contingent Valuation When Respondents Are Ambivalent." *Journal of Environmental Economics and Management* 29(2): 181-196.
- Rollins, K., D. Dumitras, and A. Castledine. 2008. "An Analysis of Congestion Effects across and within Multiple Recreation Activities." *Canadian Journal of Agricultural Economics* 56(1): 95-116.
- Rollins, K.S., M. Kobayashi, M.D.R. Evans, and A. Castledine. 2009. "Willingness to Pay Estimation When Protest Beliefs Are Inseparable from the Public Good Definition." Working Paper, Department of Resource Economics, University of Nevada, Reno. A copy of this paper is on file with the authors.
- Samnaliev, M., T.H. Stevens, and T. More. 2006. "A Comparison of Alternative Certainty Calibration Techniques in Contingent Valuation." *Ecological Economics* 57(3): 507-519.
- Shaikh, S.L., L. Sun, and G.C. van Kooten. 2007. "Treating Respondent Uncertainty in Contingent Valuation: A Comparison of Empirical Treatments." *Ecological Economics* 62(1): 115-125.
- Schwarz, N., and S. Sudman. 1995. *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. New York: Jossey Bass.
- Tourangeau, R., L.J. Rips, and K.A. Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
- Vossler, C.A., G.L. Poe, M.P. Welsh, and R.G. Ethier. 2004. "Bid Design Effects in Multiple Bounded Discrete Choice Contingent Valuation." *Environmental and Resource Economics* 29(4): 401-418.
- Wang, H. 1997. "Treatment of 'Don't Know' Responses in Contingent Valuation Surveys: A Random Valuation Model." *Journal of Environmental Economics and Management* 32(2): 219-232.
- Welsh, M., and G. Poe. 1998. "Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach." *Journal of Environmental Economics and Management* 36(2): 170-185.
- Whitehead, J., G. Blomquist, R. Ready, and J. Huang. 1998. "Construct Validity and Polychotomous Choice Contingent Valuation Questions." *Environmental and Resource Economics* 11(1): 107-116.

## Appendix 1. CV Policy Proposals

### Proposal Version A (PL)

#### PROPOSAL

**Suppose** that experts predict the numbers of wildfires in Nevada to **double** over the next five years due to the continued spread of cheatgrass. This will lead to the loss of native grasses, wildflowers, and shrubs.

A new, intensive Rangeland Vegetation Management Program has been proposed. This program will reduce fire risk by reducing cheatgrass through the use of prescribed fires, machinery, herbicides, prescribed grazing, and seeding with native plants and nonnative grasses such as crested wheatgrass.

Under this new program, fire risk would **not double**, but stay the same as it is now.

### Proposal Version B (OG)

#### PROPOSAL

**Suppose** that a new, intensive Rangeland Vegetation Management Program has been proposed. This program will reduce fire risk by reducing cheatgrass through the use of prescribed fires, machinery, herbicides, prescribed grazing, and seeding with native plants and nonnative grasses such as crested wheatgrass.

The new program could **reduce** the number of wildfires throughout the state **by half**.

## Appendix 2. Bid Presentation (M-version)

**9. Would you vote for this proposal if passage of the proposal would cost you these amounts every year for the foreseeable future?**

Please check one box for each amount.

Cost to you per year	How would you vote?				
	Definitely No	Probably No	Probably Yes	Definitely Yes	Not Sure
\$ 0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 31	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 52	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 83	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 114	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 157	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
\$ 282	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>