VALUING ACCESS TO U.S. PUBLIC LANDS: A PRICING EXPERIMENT TO INFORM FEDERAL POLICY

July 2009

David Aadland
Bistra Anatchkova
Burke Grandjean
Jason F. Shogren
Benjamin Simon
Patricia A. Taylor [†]

Abstract. We report on a nation-wide experiment to price access to U.S. public lands. The U.S. Federal Lands Recreation Enhancement Act mandated a new annual pass to cover all federal recreation sites that charge an entrance or access fee. Our task was to assist federal land management agencies in determining an appropriate price for the new pass. We administered a national telephone survey to over 3,700 households and used contingent valuation to estimate households' willingness to pay (WTP) for the pass at different prices. Our model controls for methodological problems such as starting-point bias, and our experimental design allows us to estimate the degree of hypothetical bias. We develop a novel method to convert that estimate into dollars, and thereby to calibrate our WTP results against actual purchasing decisions. In a sample of the general U.S. population – most of whom have little experience with similar federal passes – respondents greatly exaggerate their WTP for the pass when contrasted with their recent pass purchasing behavior. A sample of former pass purchasers, however, exhibits less exaggeration of WTP, confirming other research showing that market experience can mitigate hypothetical bias. We briefly describe how our conclusions were put into practice by policymakers. Finally, we demonstrate the external validity of our results by comparing the projected pass sales from our model to actual pass sales for the first two years of the new program.

[†]Author affiliations: Aadland (aadland@uwyo.edu) is associate professor of economics at the University of Wyoming, Laramie WY 82071. Anatchkova is survey manager of the Wyoming Survey & Analysis Center (WYSAC) at the University of Wyoming. Grandjean is executive director of WYSAC and professor of statistics and sociology at the University of Wyoming. Shogren is the Stroock Professor of Natural Resource Conservation and Management at the University of Wyoming. Simon is the acting assistant director for the Office of Policy Analysis, U.S. Department of the Interior. Taylor is professor of sociology at the University of Wyoming. This project was supported through a cooperative agreement between the University of Wyoming and the U.S. Department of the Interior. This paper does not reflect the official position of the University of Wyoming or the U.S. Department of the Interior. We thank Vic Adamowicz, Sharon Lohr, Kim Magraw, Bruce Peacock, Kerry Smith, as well as conference participants at the Western Economic Association and American Agricultural Economic Association annual meetings, for their valuable input.

1. Introduction

This paper discusses a nation-wide experiment that was used by federal land management agencies in pricing recreational access to U.S. public lands. The Federal Lands Recreation Enhancement Act of 2004 mandated a new interagency pass (IAP) to replace the Golden Eagle Passport (GEP). The IAP, similar to the GEP, provides the option of purchasing a single annual pass instead of paying at the gate for each visit. And like the GEP, the IAP covers all federal lands that charge an entrance or access fee for recreational use. This includes many sites administered by the U.S. Forest Service in the Department of Agriculture, units of the National Park Service in the Department of the Interior, and some locations administered by other Interior Department offices.

Our interdisciplinary research team undertook a comprehensive study of pricing for the proposed new pass and submitted our results to the agencies in the fall of 2006. After reviewing our report, the Departments of Agriculture and the Interior set the price for the IAP at $80 per year; the new pass went on sale in January 2007.

Here we describe our econometric analysis of the national survey we conducted to assist the federal agencies in pricing the pass. We used contingent valuation (CV) methods to provide that assistance. CV and other stated-preference valuation methods have made significant in-roads into public decision-making over the past two decades. Policymakers use the value estimates to help guide their decisions for a range of topics like water quality protection, air quality improvements, watershed and ecosystem protection, and reduced human health risk (see Brown *et al.*, 2004).

However, a major criticism of CV is that households tend to exaggerate their maximum willingness to pay (WTP) on a survey – the issue of *hypothetical bias*. In the present application,

we use an experimental design that allows us to contrast the hypothetical purchasing decisions of survey respondents with the actual purchasing decisions of households. Unlike many other valuation exercises,[1] we have a real benchmark – the GEP. The GEP entitled households to basically identical access to public lands as the IAP and had been offered at $65 since 2000. Using the reported GEP purchasing decisions of households within our sample, we are able to estimate the degree of hypothetical bias. We develop a novel method of expressing that bias in dollars. This makes it possible to recalibrate our price estimates for greater consistency with actual purchasing decisions.

Our survey design also allows us to compare results for the U.S. general population (in a Random Digit Dialing sample) to those for former pass purchasers (in a sample obtained from the National Parks Foundation). The findings support previous research showing that market experience is an important indicator of people's ability to state their true valuation for public and environmental goods (e.g., Cameron and Englin, 1997; List and Shogren, 1998; List and Gallet, 2001; Cherry *et al*., 2003; Murphy *et al*., 2005; and Cherry and Shogren, 2007).

For the present paper, we then use aggregate data on sales of the IAP – information that was, of course, unavailable to us or to public land managers at the time the policy decision had to be made. This new information about subsequent sales confirms the external validity of our recalibrated demand forecasts.

As an instance of applied, real-world policy research, our project had to balance the need for methodological rigor with the importance of clear communication. Our analysis required complex econometrics to generate sound conclusions. But we sought to convey those conclusions in an uncomplicated way to federal policymakers with widely varying levels of statistical expertise. The present paper reflects a similar balancing effort. We discuss issues of

policy, economics, and methodology in a manner intended to be accessible to readers across a variety of disciplines and with a range of statistical expertise. The more technical econometric details are placed in an appendix for interested readers.

2. Policy Context

The Federal Lands Recreation Enhancement Act (Public Law 108-447, the FLREA) was signed into law on December 8, 2004. It mandated a new recreation pass to replace the existing GEP and National Parks Pass (NPP). (For a list of all acronyms used in this paper, see Table 1.) The FLREA called for a pass to provide convenient and fairly priced access to federal recreation sites that charge fees; to provide opportunities for education on federal lands; to provide support for public lands; and to develop partnerships with organizations that support public lands.

In fulfilling their statutory duties under the FLREA, the Departments of Agriculture and the Interior sought our help for pricing the new pass. Their stated objective was a pass price that would result in no loss of total revenue relative to what would be generated from gate receipts alone, in the absence of any pass. This objective should be recognized as a political constraint on the price-setting exercise. There are other possible policy issues associated with pass pricing (e.g., public education, congestion, pollution, effects of visitors on infrastructure, etc.) that were outside the scope of our study. We explored neither the optimal price based on the full social costs, nor the distribution of total revenues across federal sites or agencies.

As specified in the FLREA, the new interagency pass (the IAP, also called the "America the Beautiful" pass) covers the entrance fees for units administered by the National Park Service and the U.S. Fish and Wildlife Service, as well as standard amenity fees for developed areas administered by the Bureau of Land Management, the Bureau of Reclamation, and the U.S.

Forest Service (see www.nps.gov/fees_passes.htm). Fee areas generally include visitor centers or other developments with certain basic facilities (e.g., designated parking, restrooms, etc.). Like the passes it replaced, the IAP does not cover any expanded amenity fees for such things as reservation services or developed campgrounds and boat launches. When appropriate, pass holders also continue to pay special recreation permit fees such as those for motorized recreational vehicle use, recreational events, and group activities.

The IAP is very similar to the GEP it replaced. The GEP was introduced in 1965, at a price of $7 per year; by 2000, its price had risen to $65. Like the IAP, the GEP provided entrance for the pass holder (and other occupants of a private vehicle) to virtually all federal lands that charge a fee.[2] The year 2000 saw the introduction of the NPP, sold for $50 and providing entrance only to those federal lands administered by the National Park Service. For an additional $15, a sticker could be purchased and affixed to the NPP to make it equivalent in all respects to the GEP (and therefore, essentially equivalent to the IAP). Throughout this paper, "GEP" refers to either the Golden Eagle Passport or the National Parks Pass with sticker, while "NPP" refers to the National Parks Pass without the sticker. Sales of the GEP and the NPP were discontinued in January of 2007 when the IAP was introduced.

3. Economic Approach

A purchaser of the IAP derives value from both use and non-use sources. The IAP provides access to recreation at federal sites, without paying entrance or usage fees at the gate. We refer to this as *use value*, which can in turn be separated into *convenience* and *economic value*. Convenience refers to the reduced transaction costs associated with using the IAP rather than having to make separate payments for each entrance fee.[3] Economic value is derived from the

expected cost in entrance fees, given the number of planned visits to recreation sites and the current entrance fee structure. The IAP may also provide households with a sense of satisfaction in helping to fund the creation and maintenance of federal recreation sites that may be used by others. We refer to this non-use value as *stewardship*.

Our treatment of the economic value of the pass assumes households base their purchasing decision on the expected number of recreation trips and the current gate fees, not on the amount of uncertainty associated with future trips. We recognize, however, that households might incorporate the option value of the pass into their decision-making. Similar to a *real option*, the IAP can be thought of as an irreversible investment under uncertainty (Dixit and Pindyck, 1994). Once the IAP is purchased, the decision is final and irreversible because of legal restrictions on the resale of purchased passes. Now suppose a household, having purchased the pass, receives an adverse shock (e.g., sudden loss in income, rising gasoline prices, family illness, etc.) and decides it must forgo a previously planned family vacation to federal lands. The household may *ex post* regret the decision to purchase the IAP, as the purchase no longer satisfies the internal cost-benefit test. But it also seems reasonable that the household will, at least to some degree, anticipate this possibility *ex ante* and place a value on the option to delay the purchase. While we do not formally model the option value for the IAP, we suggest the household will likely incorporate the value of delaying purchase into its stated WTP for the pass.

In deciding whether to purchase the IAP, households weigh all the benefits and costs. The benefits include convenience, economic value, and possible stewardship motives. The relevant cost is the price of the pass. The household's decision is straightforward – if the total benefits outweigh the private costs, purchase the pass; otherwise, pay the gate fees. Households with a maximum WTP above the price of the IAP will purchase the pass, while households with a

maximum WTP that is below the price of the IAP will instead pay at the gate. To better understand this decision, we undertook a CV experiment designed to simulate the actual market environment and IAP purchasing decisions.

4. Survey Design

Our goal with the survey was to link the IAP valuation exercise to real choices and real outside options. Respondents answered the new valuation question after thinking about the *status quo* and their familiarity with the NPP and GEP. Our intent was to keep the survey realistic, simple, and quick while still generating reasonable value estimates. Given time constraints and the scale of the project, a face-to-face survey was infeasible. Instead, we conducted a nationwide telephone survey between February and April of 2006. Specifics on the survey methodology (development, pre-testing, sampling, weighting, and the *verbatim* questionnaire interviewing script, with frequency distributions on all items) are publically available (see www.doi.gov/initiatives/rec_fees/NRP Survey Final.pdf).

4.1 Sampling

The survey sample consisted of two independent strata. The first stratum was sampled using a national Random Digit Dialing (RDD) draw of households with landline telephones. To enhance the precision of estimates from this part of the sample, it was pre-stratified by the 9 major geographic divisions defined by the U.S. Bureau of the Census, and then weighted to reflect Census Bureau distributions by geographic division, household income, Hispanic origin, and racial identification. By design, this sample of 1,799 responding RDD households represents

the roughly 110 million households in the U.S. with landline phones, except for about 30,000 that fall into the second stratum.[4]

The second stratum was randomly sampled from a population list of telephone numbers for households known to the National Parks Foundation (NPF) to have purchased a National Parks Pass (with or without a sticker) between April 2004 and March 2005. This sample was also pre-stratified by geographic division, and the 1,974 responding NPF households were weighted to match the geographic distribution of the NPF population list.

The NPF list mainly covers households that purchased the NPP on-line, one to two years prior to the survey. Fortunately, our interest in the NPF sample does not require that it represent all recent pass purchasers. The RDD sample is designed to be nationally representative, and therefore to include recent pass holders in proportion to their share of all households. However, that share is small, and therefore the RDD sample does not provide a sufficient number of recent pass purchasers for detailed analysis. The virtue of the NPF sample is that it ensures an ample number of respondents likely to be familiar with the passes. Analyses comparing the two strata therefore provide an opportunity to examine the relationship between market familiarity and hypothetical bias.

The total sample size across both strata was 3,773 households.[5] Both samples were screened, through questions asked early in the survey, to eliminate households in which anyone qualified for a Golden Age or Golden Access Passport. These lifetime passes, available only to senior citizens (for a $10, one-time charge) and the disabled (for free), permanently remove a household from the market for an annual GEP, NPP, or IAP. As stipulated by the sponsoring agencies, households that had not visited any federal recreation lands in the past two years were also screened out of the full survey, on the assumption they would not be part of the relevant market

either. Protest households (those that said they would not accept the IAP for free and also rejected the premise that the federal government should sell access to public lands) were excluded from our analysis as well. After all screening, the combined sample to be analyzed consisted of 2,009 cases (523 RDD and 1,486 NPF).[6]

4.2 Questionnaire

The survey questionnaire began with an introduction, screening questions, and a short series of items about the household's recent recreational experiences on federal lands. For households that met the screening criteria, the interviewer then asked whether the respondent was aware of the NPP or GEP. If so, the respondent was asked whether the household had purchased either pass in the past 12 months. The next section of the questionnaire provided a short description of the IAP followed by the valuation question, which asked whether the household would be willing to buy the IAP at a randomly selected bid value. For some respondents, the random bid value was the current GEP price, which allows us to compare hypothetical purchasing decisions to real purchasing decisions. To more precisely pinpoint the valuation distribution, the interviewer then asked a follow-up question. If the respondent was willing to purchase the IAP at the first bid price, we increased the price by a randomly set amount and asked again; if not, we lowered the price before asking again. The final section of the questionnaire asked a series of socio-demographic questions including age, education, race, and household income.

Table 2 presents the full set of survey variables used here, with definitions and descriptive statistics. Table 3 shows the percent of the RDD and NPF samples that respond "YES" to the first bid. As expected, the percent responding "YES" generally declines as the bids increase in value (with minor exceptions at the highest bids for the RDD sample). The NPF

respondents are more likely to say "YES" at every bid price, consistent with their status as recent pass purchasers. We turn now to econometric modeling of the data.

5. Econometric Analysis

We break the econometric analysis into four subsections. In Section 5.1, we present the two models used to estimate overall WTP for the IAP. We use an interval regression approach that follows directly from our double-bounded dichotomous-choice (DBDC) survey design (see Hanemann *et al.*, 1991; Wooldridge, 2002). Section 5.2 discusses the issue of hypothetical bias in relation to prior experimental and econometric literature. Section 5.3 develops the probit model used to estimate the degree of hypothetical bias in terms of increased probability that a stated-preference household will purchase the pass at $65. In Section 5.4 we describe how to translate the probability measure into a dollar-denominated measure of hypothetical bias. This measure is then used to recalibrate the WTP estimates for greater consistency with the actual purchasing decisions of households. A detailed description of the econometric analysis can be found in the accompanying Technical Appendix (available upon request from the authors).

5.1 WTP Models

We start by writing the empirical model in terms of a household's maximum WTP for the IAP, which is indirectly derived from the utility of visiting federal recreation sites.[7] We refer to this as the "initial" model (to distinguish it from the "revised" model that will be discussed shortly):

$$WTP_{1i} = X_i'\beta + \varepsilon_i, \tag{1}$$

where WTP_{1i} is the initial latent willingness to pay for the IAP; X_i is a vector of explanatory variables; β is a vector of coefficients; ε_i is a mean-zero Gaussian error term with variance σ^2; and $i = 1,..,N$ indexes households in a sample of size N. Given expression (1), the probability of purchasing the pass is represented as

$$P_i = \Pr(y_i = 1) = \Pr(WTP_{1i} \geq b_i), \tag{2}$$

where $y_i = 1$ if the household purchased the pass, $y_i = 0$ if the household did not purchase the pass, and b_i is the proposed price of the IAP (the bid). The first bid is chosen at random from the following bid vector (which we developed using focus groups):

$$\text{b} = (\$25, 45, 65, 85, 105, 125, 145, 165). \tag{3}$$

We follow the first bid with a second one, randomly selected from either b^H = (b+\$20, 2b-\$5) if the respondent accepts the first bid or b^L = (b-\$20, 0.5(b+\$5)) if the respondent declines.[8] If a respondent declines both bids, we ask another follow-up question with a bid equal to *zero*. If the respondent even says "NO" to a free pass, we ask why, in order to identify protest bidders.

The DBDC format just described was first introduced by Carson *et al*. (1986) to increase the information obtained from a traditional single-bid referendum. Hanemann *et al*. (1991) show the second bid improves the efficiency of the resulting estimates (cf. Carson, 2007). However, researchers have also questioned whether responses to the second bid are consistent with responses to the first bid. Cameron and Quiggin (1994), Herriges and Shogren (1996), Alberini *et al*. (1997), Green *et al*. (1998), and others have designed differing approaches to handle the fact that WTP may be anchored to the amount of the first bid, and/or may be shifted in

consequence of receiving a second bid. More recently, Whitehead (2002) has developed an integrated approach to test and control for both anchoring and structural shifts in multiple valuation questions.

Following Whitehead, we specify the following "revised" WTP function:

$$WTP_{2i} = (1-\gamma)WTP_{1i} + \gamma b_i + \lambda, \quad (4)$$

where the revised WTP is a weighted average of initial WTP and the bid amount, plus a shift factor. The parameter γ captures the anchoring effect ($0 \leq \gamma \leq 1$) and λ captures possible structural shifts, which may include nay-saying, yea-saying, or other respondent reactions to being presented with a second bid (Kanninen, 1995). The parameter estimates, obtained through maximum likelihood methods, are used to form WTP estimates for every household in the sample.[9]

As a preliminary step in our modeling, we test whether the RDD and NPF samples can be pooled together. A likelihood ratio test soundly rejects the hypothesis that the β vectors are equal across the two populations [$\chi^2(21) = 474.30, p < 0.01$]. We then test the RDD and NPF coefficients individually and find that we cannot reject the equality of several pairs of coefficients. In each instance, coefficients that do not differ statistically are also substantively similar. Therefore, we pool the RDD and NPF samples, but we allow coefficients to vary across the two samples if they are significantly different. Coefficients for which the difference between the RDD and NPF samples does not approach statistical significance ($p > 0.1$) are restricted to be the same. A joint likelihood ratio test of the individual parameter restrictions also supports imposing equal coefficients ($LR = 17.54,\ p > 0.1$).

Table 4 presents the results from the pooled samples for both the standard DBDC model and Whitehead's model that adjusts for the effects of a second bid. We include a number of different control variables, several of which have consistently significant and theoretically plausible effects across the two samples and two models. For example, those with lower WTP for the pass include residents of the Western South Central region (i.e., the states of AL, KY, MS and TN) and the Great Plains (IA, KS, MN, MO, ND, NE, SD). These are areas with comparatively few federal lands; hence for their residents, consideration of the travel cost to recreate on federal lands may far outweigh the total convenience, economic, and stewardship value of a pass. Conversely, residents of the Rocky Mountain region (AZ, CO, ID, MT, NM, NV, UT and WY) have the highest WTP. This region includes many large national parks, national forests, and other federal lands in relatively close proximity to population centers. Much the same is true of states on the Pacific Coast (AK, CA, HI, OR, WA), the reference category in the set of region variables. The Rocky Mountain area is the only region with a significantly higher WTP than the Pacific Coast. While these results are suggestive, we leave detailed examination of the effects of distance, and the related issue of travel costs, as possible topics for future research.[10]

In both samples, low income households have lower WTP, as expected, but this effect is only significant among the NPF respondents. Households preferring that pass revenue be explicitly directed toward facilities and services within the recreational sites have a significantly higher WTP in both samples, perhaps indicating a stewardship motive. In both samples, a spline function fits the relationship between past number of visits to recreation sites and WTP for the pass. For households with four or fewer past trips, WTP increases as the number of trips increases, as expected if past use is positively correlated with expected future use. The relationship flattens out beyond four trips, so that additional trips do not lead to higher WTP.

Only the NPF sample exhibits significant anchoring, with a higher starting bid associated with a higher reported WTP. Both the RDD and NPF samples show evidence of a negative shift in WTP between bids. Having stated a decision on the first bid, respondents are less likely to accept the second bid (all else equal) than they would have been had they been offered that same bid amount first. The significant shift coefficient controls for a downward bias in overall WTP estimates that would otherwise result from the DBDC design (Whitehead, 2002). As expected (see Technical Appendix), the model is significantly heteroscedastic, with the variance in WTP rising as the bid value increases. Error variance is higher in the RDD sample than among NPF households, but that difference drops to insignificance when anchoring and shift are controlled.

Model-derived mean and median WTP figures are presented at the bottom of Table 4. Using the Whitehead model, the mean and median WTP values for the RDD sample are approximately $33 and $29, indicating a WTP distribution that is slightly skewed toward higher WTP values. The mean and median WTP values for the NPF sample are approximately $83 and $76 under that model, again indicating a distribution skewed to the right. The mean and median WTPs are about two and one-half times higher in the NPF sample than in the RDD sample, which is expected given that the NPF sample is comprised of former pass purchasers. The means and medians based on the Whitehead model are also higher than those based on the standard DBDC model, because the latter does not control for the downward shift in WTP on the second bid.

5.2 Prior Research on Hypothetical Bias

A fundamental concern of any CV study is *hypothetical bias*. Respondents have a well-established tendency to state willingness to pay values that are significantly greater than those revealed in real-market interactions.

In his seminal experimental work, Bohm (1972) compared bids in hypothetical and actual (albeit experimental) markets that elicited subjects' stated value to preview a Swedish television show. His results suggested that people overstate their actual values when asked a hypothetical question. This conclusion was further buttressed by Bishop and Heberlein's (1979) pioneering work eliciting real and hypothetical values for goose and deer hunting permits in Wisconsin. Subsequent research has generally supported the observation that hypothetical values exceed actual WTP (e.g., Seip and Strand, 1992; Neill *et al*., 1994; Frykblom, 1997; Blumenschein *et al*., 2008).[11] Most evidence suggests the average person exaggerates his or her actual WTP across a broad spectrum of goods with vastly different experimental parameters (e.g., see Harrison and Rutström's 2006 review of values elicited for about forty private and public goods). Research indicates that the hypothetical-to-actual WTP ratio can range from -43% to +2600% for goods such as irradiated pork, watercolor paintings, and maps (see also Diamond and Hausman's review, 1994). This body of evidence reinforces the key point that a gap exists between stated intentions and actual behavior, triggering a large literature in valuation research to understand why the gap exists and what can be done to minimize it (e.g., see Murphy *et al*., 2005).

Several different methods have been proposed to obtain more accurate WTP values from valuation surveys (Boyle, 2003). The approaches can be grouped into two broad categories: *ex ante* and *ex post*. *Ex ante* methods involve what Cummings *et al.* (1997) call "instrument calibration," whereby the questionnaire is crafted to minimize hypothetical bias. Such methods include framing the valuation question so as to make it more "real" to the respondent, providing additional "cheap talk" information to get people to realize that they might be overstating their actual preferences, or asking a follow-up question about the certainty of a person's stated valuation (e.g., Cummings and Taylor, 1999; Blumenschein *et al*., 2008). *Ex post* techniques, or

"statistical calibrations" (Cummings *et al.*, 1997), attempt to adjust for hypothetical bias by calibrating stated values to price data from real market proxies (e.g., Champ *et al*., 1997; Blackburn *et al*., 1994; Fox *et al*., 1998).

Herein we use both an *ex ante* and *ex post* techniques to mitigate hypothetical bias. We employ an *ex ante* method in the design of our survey, reminding respondents as part of the bid question that "anything you spend on the new pass will reduce the money available for your household's other expenses. And of course, there are other places for recreation, such as state parks, local historic sites, and private land." In addition, we use an *ex post* method for calibrating households' bids against actual purchasing of the (now defunct) GEP, as described next.

5.3 Estimating Hypothetical Bias as a Probability

We present a two-stage econometric procedure to calibrate hypothetical WTP values to be consistent with real choices. The first stage, discussed in this section, estimates a probit model to quantify the difference between the hypothetical decisions to buy the IAP (at a bid equal to $65) versus actual decisions to purchase the then-existing, $65 GEP.[12] The result is a probability estimate for the difference between hypothetical and actual purchase of the IAP at $65. In the second stage (Section 5.4), we will use the estimation from the DBDC model and the entire WTP distribution to back out a dollar-denominated WTP calibration factor consistent with the estimated probability difference obtained in this first stage.

Our comparison of hypothetical and actual decisions at a single price is similar to the analysis of Aadland and Caplan (2003) for recycling services; Cameron *et al*. (2002) for a "green" program to use renewable energy and plant trees; and Cummings *et al*. (1997) for a

"citizens guide" to inform residents of potential groundwater contamination. We first specify a probit model with a dummy variable to capture the difference between hypothetical and actual purchasing decisions. Within the pooled RDD/NPF sample, we analyze data from two distinct subsamples—the *stated preference* and the *revealed preference* subsamples. The revealed preference (RP) subsample includes every household that knew of the $65 GEP, whether or not they had purchased one. The stated preference (SP) subsample includes every household that was unaware of the GEP and received a bid of $65 for the IAP, whether or not they said "YES" to that bid.[13]

The underlying economic model is similar to (1) but modified to incorporate hypothetical bias:[14]

$$WTP_i = X_i'\beta + \delta SP_i + \varepsilon_i \tag{5}$$

where WTP_i is unobserved willingness to pay for either the GEP or the IAP, SP_i is a dummy variable equal to one for the stated-preference subsample and zero for the revealed-preference subsample, and δ is the hypothetical bias coefficient.

In standard probit models with fixed cut points, the coefficients are only identifiable up to a scale factor. However, the *marginal effects*, which measure the change in probability for a unit change in the explanatory variables (X or SP), are identifiable. We are primarily interested in the marginal effect for hypothetical bias:

$$\Delta = \Pr(y = 1|X, SP = 1) - \Pr(y = 1|X, SP = 0) \tag{6}$$

Before turning to our empirical estimation of hypothetical bias, we note that there are at least two reasons to expect the estimate to be a lower bound on the true total bias in the population. First, Murphy *et al*. (2005) report that hypothetical bias increases with bid level, which would mean our estimate is too small at prices higher than $65. Because policymakers ultimately decided to price the IAP at $80, we may be underestimating the hypothetical bias at that price. Second, our estimation is based on partitioning the sample according to whether respondents were aware of the GEP (revealed-preference households) or not aware of the GEP (stated-preference households that received a $65 bid for the IAP). If lack of awareness of the GEP is associated with a lower salience of federal recreation opportunities (and therefore a lower WTP for the IAP), then our stated-preference households are, all else equal, less likely to hypothetically purchase a pass at any given price. This is therefore a second reason to expect our estimate of hypothetical bias to be conservative. Nevertheless, even this conservative adjustment should substantially improve the consistency of our CV results with actual purchasing decisions.

As before, we begin the empirical exploration of the model by testing whether the RDD and NPF samples can appropriately be pooled.[15] A likelihood ratio test fails to reject the hypothesis that all corresponding elements of the β vectors are equal between the two populations $[\chi^2(10) = 13.08, p > 0.21]$. The coefficients are also substantively similar. Therefore, we pool the two samples and constrain all coefficients to be the same other than the hypothetical bias coefficient.

The first two rows of Table 5 present the coefficient and marginal effect (ME) estimates for hypothetical bias. In both samples, the hypothetical bias coefficients are positive and statistically significant. However, the coefficient and associated ME are much larger in the RDD sample. The ME estimates indicate that, all else equal, the average RDD and NPF stated-preference

households are 19.6 and 11.3 percentage points more likely to purchase a hypothetical $65 IAP than similar revealed-preference households are to have purchased the actual $65 GEP. The control variables include demographics and a region dummy. Consistent with the full DBDC models of Section 5.1, residents in the Rocky Mountain and Pacific Coast states are more likely to purchase a pass, and the positive effect of frequent prior visitation again flattens out beyond four visits to federal lands.

5.4 Transforming Hypothetical Bias from Probability to Dollars

The coefficient of hypothetical bias, Δ, in expression (6) is measured in terms of probability the pass is purchased. Although Table 5 provides evidence of significant hypothetical bias, it does not allow household WTP (measured in dollars) to be directly adjusted to reflect revealed preferences. Furthermore, the probit model in (5) does not identify a dollar amount of hypothetical bias because the bids are not varied (i.e., the bid is fixed at $65). Fortunately, the results reported in Section 5.1 are based on the complete bid vector and thus allow us to identify the entire distribution of household WTP. To the best of our knowledge, this is the first attempt at transforming a probability measure of hypothetical bias (resulting from a stated vs. revealed preference comparison at a single price) into a dollar-denominated WTP calibration factor (using a DBDC estimate of the entire WTP distribution).[16]

Figure 1a illustrates the procedure for identifying δ, the WTP calibration factor for hypothetical bias for the RDD sample, and Figure 1b illustrates the procedure using the NPF sample. The figures are drawn separately to reflect the position and shape of the empirical WTP distributions for each sample. The etched region in Figure 1a is Δ, which is the difference between the area under the *hypothetical* WTP distribution to the right of $65 and the area under

the *actual* WTP distribution to the right of $65. (In Figure 1b, Δ is the difference between etched areas H and A.) Using this probabilistic measure of the bias, we then infer the value of δ that positions the hypothetical WTP distribution to be consistent with Δ and the standard deviation σ. The corresponding WTP calibration factor, δ, is then used to form the *calibrated* WTP estimates ($WTP_i - \delta$) that more accurately reflect actual purchasing decisions of households. Table 6 reports the details for this calculation.

The estimated value for the hypothetical bias calibration factor, δ, is approximately $26 and $16 for the RDD and NPF samples, respectively. Put differently, the dollar values for mean WTP in the RDD and NPF samples need to be reduced by approximately 80% and 20% (see the last row of Table 6) to be consistent with the observation (from Table 5) that stated-preference households are 19.6 and 11.3 percentage points more likely than revealed-preference households to purchase a pass at $65. When estimating the demand curve for sales of the IAP, δ should be used as an adjustment factor to reflect actual purchasing decisions more accurately.

6. Policy Application and Outcome Check

The analytical methods discussed in this paper have been refined and elaborated since we submitted our report to the sponsoring federal agencies (www.doi.gov/initiatives/rec_fees/NRP Economic Report.pdf). For example, our econometric model has been re-specified to control explicitly for anchoring and structural shift under DBDC. However, the analyses presented here yield similar conclusions to the original, time-urgent report to policymakers.

After calibrating the WTP estimates for hypothetical bias, we submitted our pricing analysis for the IAP in the fall of 2006. We concluded that a price in excess of $100 would be necessary to achieve the agencies' original objective of preventing revenue loss from the households that

would have paid more at the gate than the price of the pass. (The present results would lead to the same conclusion, given the same policy objective.) We also provided forecasts of the expected revenue loss at lower prices, which we estimated to be modest at pass prices of $75 or above. After receiving our results and conclusions, the federal agencies officially announced the new recreation pass in December 2006 and set the price at $80. Ultimately, policymakers decided that some forgone revenue was preferable to the adverse public reaction that would likely have greeted a new pass priced at close to double the National Parks Pass or Golden Eagle Passport.

The new pass went on sale in January, 2007. The Department of the Interior has now provided us with figures on pass sales during the first two years of the program.[17] This gives us a rare opportunity for an external validity check on our model.

As shown in Table 7, actual sales of the IAP in 2007 totaled about 340,000 for the year.[18] With no calibration for hypothetical bias, the demand curve derived from our econometric analysis of the CV survey data would project sales of 1,470,000 passes at a price of $80.[19] This is almost five times too high and reflects the well-established tendency of survey respondents to exaggerate their willingness to purchase goods in hypothetical markets. However, after adjusting the WTP estimates using our hypothetical bias calibration factors ($26 for RDD and $16 for NPF), we project annual sales to be approximately 370,000.

Thus, in the first year of the new pass, actual sales were within 10% of the calibrated projection. We somewhat overestimated demand, as expected, because we used a lower-bound estimate of hypothetical bias. In the second year of the IAP, sales fell still farther below our projection. However, low sales of the $80 IAP in 2008 may well have been due to a general weakening of the economy and/or to gasoline prices reaching record highs, both of which would

likely reduce visits to federal lands. Consistent with this interpretation, sales of the $10 Senior pass also declined in 2008.

In sum, the results of our external validity check demonstrate that a carefully designed use of revealed and stated preference data can generate adjusted WTP estimates of sufficient accuracy to provide valuable guidance to policymakers on the likely consequences of their decisions.

7. Concluding Remarks

Stated preference valuation methods have made significant in-roads into public decision-making over the past two decades. Policymakers use the value estimates to help guide their decisions for a range of topics like water quality protection, air quality improvements, watershed and ecosystem protection, and reduced human health risk (see for example Brown *et al.*, 2004). Here we have described our use of contingent valuation to assist federal land management agencies to determine an appropriate price for a new pass providing access to all federal recreational lands. This project was spurred by the U.S. Federal Lands Recreation Enhancement Act of 2004, which mandated a new pass to replace the existing Golden Eagle Passport and the National Parks Pass.

Our comprehensive study involved meetings with federal policymakers and national stakeholders, focus groups with diverse citizens, interviews with officials of the Canadian parks agency and various state park systems, a nation-wide telephone survey, and an econometric analysis of the survey data. We designed and administered a contingent valuation survey to over 3,700 households in two samples to estimate willingness to pay for the new pass at a variety of prices. The econometric specification incorporated recent advances that account for issues such

as anchoring bias and structural shifts in stated WTP that are associated with double-bounded dichotomous choice models.

Our experimental design allowed us to estimate the degree of hypothetical bias in the sample. We developed a novel method to express hypothetical bias in dollars, and thereby to calibrate our WTP estimates to reflect actual purchasing decisions. We were also able to demonstrate the external validity of our model and findings by comparing results derived from the modeling to sales figures on the new pass.

We find the general U.S. population – most of whom have little experience with similar federal passes – tend to exaggerate their WTP for the pass when contrasted with previous pass sales. A sample of recent pass purchasers, however, exhibits less hypothetical bias, supporting the idea that market experience can help mitigate valuation biases. More generally, we demonstrate how a carefully designed use of CV modeling can generate WTP estimates of sufficient accuracy to provide valuable guidance to policymakers on the likely consequences of their decisions.

References

Aadland, D. and A. Caplan. 2003. "Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias." *American Journal of Agricultural Economics*, 85(2), 492–502.

Aadland, D. and A. Caplan. 2004. "Incentive incompatibility and starting-point bias in iterative valuation questions: Comment." *Land Economics*, 80(2), 312-315.

Adamowicz, W., J. Louviere and M. Williams. 1994. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." *Journal of Environmental Economics and Management*, 26, 271-292.

Alberini, A., B. Kanninen and R. Carson. 1997. "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data." *Land Economics*, 73(3), 309-324.

Azevedo, C., J. Herriges and C. Kling. 2003. "Combining revealed and stated preferences: Consistency tests and their interpretations." *American Journal of Agricultural Economics*, 85(3), 525-537.

American Association for Public Opinion Research. 2006. *Standard definitions: Final dispositions of case codes and outcome rates for surveys. www.aapor.org/uploads/standarddefs_4.pdf.*

Bishop, R. and T. Heberlein. 1979. "Measuring values of extramarket goods: Are indirect measures biased?" *American Journal of Agricultural Economics*, 61, 926-930.

Blackburn, M., G. Harrison and E. Rutström. 1994. "Statistical bias functions and informative hypothetical surveys." *American Journal of Agricultural Economics*, 76, 1084-1088.

Blumenschein, K., G. Blomquist, M. Johannesson, N. Horn, and P. Freeman. 2008. "Eliciting willingness to pay without bias: Evidence from a field experiment." *Economic Journal*, 118, 114-137.

Bohm, P. 1972. "Estimating demand for public goods: An experiment." *European Economic Review*, 3, 111-130.

Boyle, K. 2003. "Contingent valuation in practice," in P. Champ, K. Boyle and T. Brown (editors), *A Primer on Nonmarket Valuation*. Kluwer Academic Publishers.

Brown, T., P. Champ and K. Boyle. Eds. 2003. *A primer on nonmarket valuation*. Kluwer Academic Publishers.

Cameron, T. and M. James. 1987. "Efficient estimation methods for 'close-ended' contingent valuation surveys." *Review of Economics and Statistics*, 69, 269-276.

Cameron, T. and J. Englin. 1997. "Respondent experience and contingent valuation of environmental goods." *Journal of Environmental Economics and Management*, 33, 296-313.

Cameron, T. and J. Quiggin. 1994. "Estimation using contingent valuation data from a dichotomous choice with follow-up questionnaire." *Journal of Environmental Economics and Management*, 1994, 27(3), 218-234.

Cameron, T., G. Poe, R. Ethier and W. Schulze. 2002. "Alternative non-market value-elicitation methods: Are the underlying preferences the same?" *Journal of Environmental Economics and Management*, 44(3), 391-425.

Carson, R. 2007. *The stated preference approach to environmental valuation*, *volume III*. Ashgate Publishing, Aldershot, U.K.

Carson, R., W. Hanemann and R. Mitchell. 1986. "Determining the demand for public goods by simulating referendums at different tax prices." Manuscript, University of California, San Diego.

Champ, P., R. Bishop, T. Brown and D. McCollum. 1997. "Using donation mechanisms to value nonuse benefits from public goods." *Journal of Environmental Economics and Management*, 33(2), 151-162.

Cherry, T., T. Crocker and J. Shogren. 2003. "Rationality spillovers." *Journal of Environmental Economics and Management*, 45, 63-84.

Cherry, T. and J. Shogren. 2007. "Rationality crossovers." *Journal of Economics Psychology,* 28, 261-277.

Cummings, R. and L. Taylor. 1999. "Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method." *American Economic Review,* 89(3), 649-666.

Cummings, R., G. Harrison and E. Rutrström. 1995. "Homegrown values and hypothetical surveys: Is the dichotomous-choice approach incentive compatible?" *American Economic Review*, 85, 260-266.

Cummings, R., S. Elliot, G. Harrison and J. Murphy. 1997. "Are hypothetical referenda incentive compatible?" *Journal of Political Economy*, 105, 609-621.

Diamond, P. and J. Hausman. 1994. "Contingent valuation: Is some number better than none?" *Journal of Economic Perspectives*, 8(4), 45-64.

Dickie, M., A. Fisher and S. Gerking. 1987. "Market transactions and hypothetical demand data: A comparative study." *Journal of the American Statistical Association*, 82, 69-75.

Dixit, A. and R. Pindyck. 1994. *Investment under uncertainty.* Princeton University Press, Princeton, NJ.

Dorofeev, S. and P. Grant. 2007. *Statistics for real-life sample surveys: Non-simple-random samples and weighted data.* Cambridge University Press, Cambridge, UK.

Fox, J., J. Shogren, D. Hayes and J. Kliebenstein. 1998. "CVM-X: Calibrating contingent values with experimental auction markets." *American Journal of Agricultural Economics*, 80(3), 455-465.

Frykblom, P. 1997. "Hypothetical question modes and real willingness to pay," *Journal of Environmental Economics and Management*, 34, 275-287.

Green, D., K. Jacowitz, D. Kahneman and D. McFadden. 1998. "Referendum contingent valuation, anchoring, and willingness to pay for public goods." *Resource and Energy Economics*, 20(2), 85-116.

Hanemann, W., J. Loomis and B. Kanninen. 1991. "Statistical efficiency of double bounded dichotomous choice contingent valuation." *American Journal of Agricultural Economics*, 73(4), 1255-1263.

Harrison G. and E. Rutström. 2006. "Experimental evidence on the existence of hypothetical bias in value elicitation methods," in *Handbook of Results in Experimental Economics*, C. Plott and V. Smith (editors). Elsevier Science, New York.

Hensher D., J. Louviere and J. Swait. 1999. "Combining sources of preference data." *Journal of Econometrics*, 89, 197-221.

Herriges, J. and J. Shogren. 1996. "Starting point bias in dichotomous choice valuation with follow-up questioning." *Journal of Environmental Economics and Management*, 30, 112-131.

Kanninen, B. 1995. "Bias in discrete response contingent valuation." *Journal of Environmental Economics and Management*, 28, 114-125.

List, J. 2001. "Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards." *American Economic Review*, 91(4), 1498-1507.

List, J. and J. Shogren. 1998. "Calibration of the differences between actual and hypothetical valuations in a field experiment." *Journal of Economic Behavior and Organization*, 37, 193-205.

List J. and C. Gallet. 2001. "What experimental protocol influence disparities between actual and hypothetical stated values? Evidence from a meta-analysis." *Environmental and Resource Economics*, 20, 241-254.

Loomis, J., T. Brown, T. Lucero and G. Peterson. 1996. "Improving validity experiments of contingent valuation methods: Results of efforts to reduce the disparity of hypothetical and actual willingness to pay." *Land Economics*, 72, 450-461.

McConnell, K., Q. Weninger and I. Strand. 1999. "Joint estimation of contingent valuation and truncated recreational demands," in *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, J. Herriges and C. Kling (editors), Edward Elgar Publishing Inc., North Hampton, MA.

Murphy, J., P. Allen, T. Stevens and D. Weatherhead. 2005. "A meta-analysis of hypothetical bias in contingent valuation." *Environmental and Resource Economics*, 30(3), 313-325.

Neill, H., R. Cummings, P. Ganderton, G. Harrison and T. McGuckin. 1994. "Hypothetical surveys and real economic commitments." *Land Economics*, 70, 145-154.

Seip, K. and J. Strand. 1992. "Willingness to pay for environmental goods in Norway: A contingent valuation study with real payment." *Environmental and Resource Economics*, 2, 91-106.

Smith, V. and C. Mansfield. 1998. "Buying time: Real and hypothetical offers." *Journal of Environmental Economics and Management*, 36, 209-224.

Whitehead, J. 2002. "Incentive incompatibility and starting-point bias in iterative valuation questions." *Land Economics*, 78(2), 285-297.

Wooldridge, J. 2002. *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, MA.

Table 1. Acronyms Used in the Text

Acronym	Full Expression
CDF	Cumulative Distribution Function
CV	Contingent Valuation
DBDC	Double-Bounded Dichotomous Choice
IAP	Interagency Pass
FLREA	Federal Lands Recreation Enhancement Act
GEP	Golden Eagle Passport (and/or National Parks Pass with sticker)
ME	Marginal Effect
NPF	National Parks Foundation
NPP	National Parks Pass
RDD	Random Digit Dialing
RP	Revealed Preference
SBDC	Single-Bounded Dichotomous Choice
SP	Stated Preference
WTP	Willingness to Pay

Table 2. Variable Definitions and Sample Means

Variables	Definitions	Means	
		RDD Sample (N = 523)	NPF Sample (N = 1486)
First Bid	Starting bid, chosen randomly from {$25,$45,$65,$85,$105,$125,$145,$165}	82.671	96.053
All Revenue	1 if household is more likely to purchase pass if all money was used for services and facilities on federal recreation sites; 0 otherwise	0.496	0.627
Visits	Number of visits last year to Federal recreation sites	1.772	5.943
High Visits	(Visits – 4) if Visits > 4; 0 otherwise	0.402	3.058
Young	1 if 17 < Age < 30; 0 otherwise	0.136	0.138
Old	1 if 45 < Age < 60; 0 otherwise	0.367	0.493
Male	1 if Male; 0 otherwise	0.402	0.474
College	1 if taken university/community college/technical school courses, has BS degree or taken some graduate or professional courses; 0 otherwise	0.508	0.558
Professional degree	1 if Graduate or professional degree; 0 otherwise	0.212	0.351
Low Income	1 if Household income is less than $50K; 0 otherwise	0.404	0.259
Hispanic	1 if Hispanic; 0 otherwise	0.077	0.029
White	1 if White; 0 otherwise	0.704	0.867
Asian	1 if Asian; 0 otherwise	0.020	0.029
African American	1 if African American; 0 otherwise	0.047	0.003
NE	1 if states (CT, MA, ME, NH, RI, VT); NE = 0 otherwise	0.051	0.069
ENC	1 if states (NJ, NY, PA); ENC = 0 otherwise	0.110	0.125
WNC	1 if states (IL, IN, MI, OH, WI); WNC = 0 otherwise	0.151	0.133
GP	1 if states (IA, KS, MN, MO, ND, NE, SD); GP = 0 otherwise	0.068	0.059
SE	1 if states (DC, DE, FL, GA, MD, NC, SC, VA, WV); SE = 0 otherwise	0.139	0.184
WSC	1 if states (AL, KY, MS, TN); WSC = 0 otherwise	0.055	0.026
SW	1 if states (AR, LA, OK, TX); SW = 0 otherwise	0.064	0.070
RM	1 if states (AZ, CO, ID, MT, NM, NV, UT, WY); RM = 0 otherwise	0.091	0.115
PC	1 if states (AK, CA, HI, OR, WA); PC = 0 otherwise	0.174	0.215

Table 3. Percent Responding "YES" to First Bid

RDD Sample (N = 523)		NPF Sample (N = 1486)	
Price	Percent Responding "YES"	Price	Percent Responding "YES"
25	49.3%	25	87.5%
45	34.6%	45	78.1%
65	20.0%	65	58.8%
85	19.4%	85	43.9%
105	16.7%	105	33.0%
125	9.8%	125	22.8%
145	11.1%	145	22.0%
165	11.3%	165	13.5%

Table 4. WTP Estimates

Explanatory Variables[†]	Whitehead DBDC Model				Standard DBDC Model			
	RDD Sample[††]		NPF Sample		RDD Sample		NPF Sample	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
RDD	-22.02	21.87	--	--	-24.99	20.08	same	--
All Revenue	18.43***	3.89	same	--	15.18***	3.29	same	--
Visits	4.40***	1.32	same	--	3.67***	1.13	same	--
High Visits	-4.12***	1.43	same	--	-3.44***	1.23	same	--
Young	-7.81	5.72	same	--	-7.52	4.95	same	--
Old	-2.70	3.98	same	--	-2.58	3.44	same	--
Male	-1.89	3.64	same	--	-1.91	3.15	same	--
College	0.41	10.62	9.50	6.41	0.37	10.03	7.86*	5.40
Professional degree	-0.93	12.43	13.44**	6.66	-0.97	11.74	11.24**	5.61
Low Income	-5.39	8.77	-15.20***	4.19	-5.42	8.24	-13.37***	3.53
White	0.42	19.46	-0.86	7.25	0.68	18.42	-1.23	6.14
Hispanic	11.56	22.37	-2.67	12.05	13.15	21.08	-4.42	10.11
Asian	11.73	28.37	-14.64	12.18	12.66	26.54	-13.80	10.28
African American	-36.07	26.36	57.88	41.28	-34.46	24.72	58.18*	35.11
NE	1.43	7.78	same	--	0.98	6.69	same	--
ENC	-9.49*	6.37	same	--	-8.98	5.49	same	--
WNC	-5.66	6.28	same	--	-4.31	5.43	same	--
GP	-14.42*	8.28	same	--	-12.84*	7.14	same	--
SE	-5.15	5.71	same	--	-5.53	4.92	same	--
WSC	-21.79*	11.36	same	--	-16.39*	9.81	same	--
SW	-5.73	7.76	same	--	-5.12	6.71	same	--
RM	22.39*	14.53	13.68**	6.68	22.66*	13.65	12.73**	5.76
Anchoring (γ)	0.11	0.11	0.26***	0.03	--	--	--	--
Shift (λ)	-8.98***	0.55	-6.81***	0.85	--	--	--	--
Heteroscedasticity (H^d) Results								
Constant	6.81***	0.18	same	--	6.38***	0.13	same	--
First Bid	0.01***	0.00	same	--	0.02***	0.00	same	--
RDD	0.18	0.23	--	--	0.44***	0.14	--	--
Likelihood Ratio Tests								
Null Hypothesis	Statistic	df	P-Value					
H_0: $\beta = 0$	385.40***	30	0.000					
H_0: $\beta_{RDD} = \beta_{NPF}$	474.30***	21	0.000					
Summary Statistics								
Sample Size	523		1486		523		1486	
Mean WTP	$32.64		$83.38		$27.59		$75.12	
Median WTP	$29.24		$76.04		$24.21		$74.26	

Notes. (***), (**), and (*) refer to statistical significance at the 1, 5 and 10 percent levels. All significance tests on the coefficients are two-tailed. The estimation was carried out using the Constrained Maximum Likelihood (CML 2.0) package in Gauss version 8.0. The nonlinear optimization routine was Newton-Raphson with a convergence criterion of 1×10^{-5} for the gradient of the coefficients. The estimates for "don't know" and "missing" dummy variables are not shown. [†]Although not explicitly listed as an explanatory variable, the bids are incorporated through the probabilities. See Cameron and James (1987) for further details. The maximum likelihood estimation incorporates weights to ensure the samples match their respective populations. The correlation parameter, ρ, is equal to 1.0. [††]The word "same" in the NPF column means that the coefficient is restricted to be equal across the RDD and NPF samples.

Table 5. Hypothetical Bias Probit Model

Explanatory Variables	Pooled RDD – NPF Sample (N = 1045)				
	Coefficient	SE	ME	Mean	
				RDD (N = 201)	NPF (N = 844)
RDD SP	63.539***	0.272	--	0.229	--
NPF SP	64.646***	0.253	--	--	0.070
RDD RP	62.433***	0.275	--	0.771	--
NPF RP	64.323***	0.194	--	--	0.930
RDD Hypo. Bias[†]	1.106***	0.318	0.196	--	--
NPF Hypo. Bias[†]	0.323**	0.189	0.113	--	--
All Revenue	0.047	0.094	0.015	0.542	0.624
Visits	0.185***	0.033	0.056	2.124	5.637
High Visits	-0.188***	0.036	-0.001	0.408	2.751
Male	0.070	0.090	0.022	0.498	0.480
Young	0.065	0.139	0.021	0.129	0.151
Old	0.143	0.099	0.046	0.453	0.505
College	0.078	0.156	0.025	0.562	0.588
Professional degree	0.116	0.166	0.037	0.249	0.321
Low Income	-0.115	0.104	-0.037	0.308	0.268
West (RM and PC)	0.296***	0.096	0.094	0.309	0.303

Likelihood Ratio Test

Null Hypothesis[††]	Statistic	Degrees of Freedom	P – Value
H_0: $\beta = 0$	56.950***	10	0.000

Notes. (***), (**), and (*) refer to statistical significance at the 1, 5 and 10 percent levels. All significance tests on the coefficients are two-tailed except for the hypothetical bias coefficients, which are run as one-tailed tests. The estimation was carried out using the Constrained Maximum Likelihood (CML 2.0) package in Gauss version 8.0. The nonlinear optimization routine was Newton-Raphson with a convergence criterion of 1×10^{-5} for the gradient of the coefficients. The estimates for "don't know" and "missing" dummy variables are not shown. The maximum likelihood estimation incorporates weights to ensure the samples match their respective populations. SE = Standard Error. ME = Marginal Effect. [†]The estimated RDD hypothetical bias coefficient (δ/σ) is calculated as the RDD SP coefficient minus the RDD RP coefficient. The NPF hypothetical bias coefficient is calculated as the NPF SP coefficient minus the NPF RP coefficient. The constant is omitted because the four dummy variables RDD SP, NPF SP, RDD RP and NPF RP encompass the entire sample. [††]The β vector does not include the RDD Hypothetical, NPF Hypothetical or associated coefficients.

Table 6. Information Used in Solving for the Hypothetical Bias Factor

Estimate	RDD Sample	NPF Sample
Δ	0.196	0.113
σ	51.40	48.14
δ	\$25.85	\$15.82
$\bar{X}'\beta$ (Mean WTP)	\$32.64	\$83.38
$\delta/\bar{X}'\beta$	0.79	0.19

Notes. The parameter estimates are derived from the following sources: Δ from Table 5; σ and $\bar{X}'\beta$ from Table 4; and δ from equation (5).

Table 7. Model-projected vs. Actual Passes Sold

	Projected Annual Passes Sold		Approximate Actual Passes Sold	
Pass	Uncalibrated	Calibrated	FY 2007	FY 2008
IAP @ $80	1,470,000	370,000	340,000	315,000
Senior @ $10	--	--	465,000	445,000

Notes. "Actual" sales are unofficial figures obtained from the U.S. Department of the Interior. For FY 2007, IAP actual sales are extrapolated to a full year from figures for January through September, and Senior actual sales include Golden Age passes sold in the first three months of that fiscal year. Model-based projections exclude both Senior and Golden Age passes.

Figure 1a. Illustration of the Procedure to Identify Dollar-Valued Hypothetical Bias (RDD Sample)

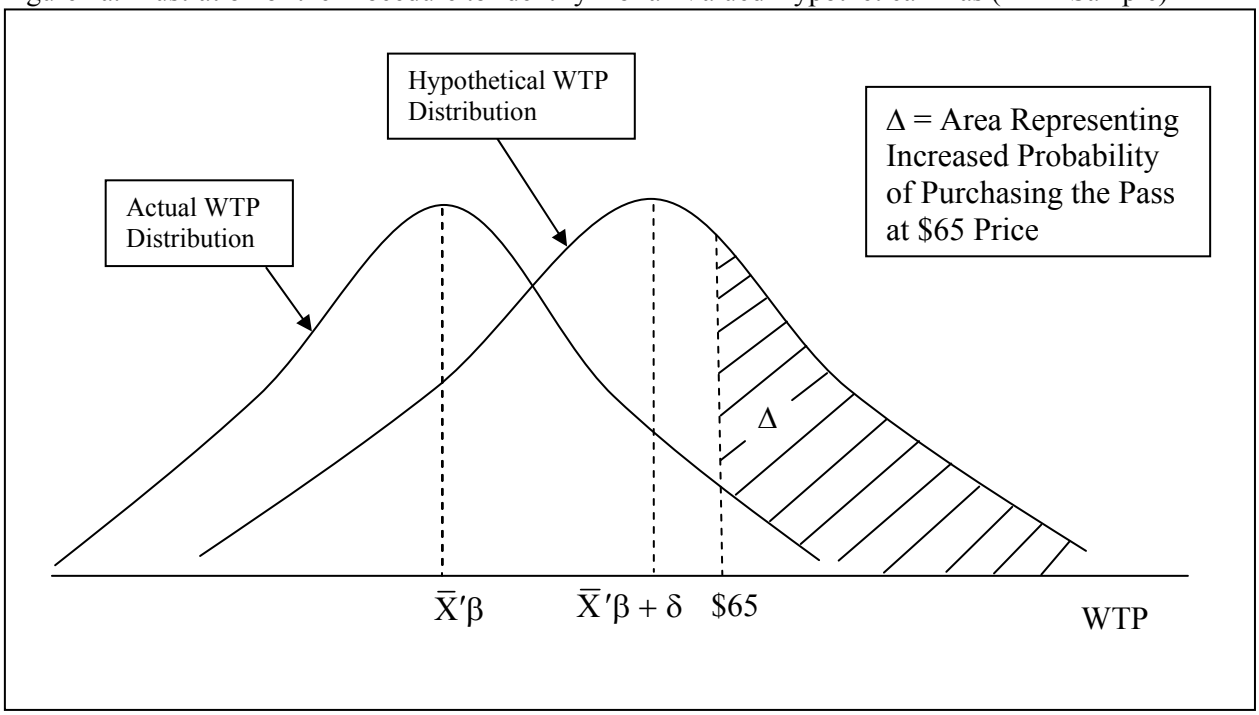
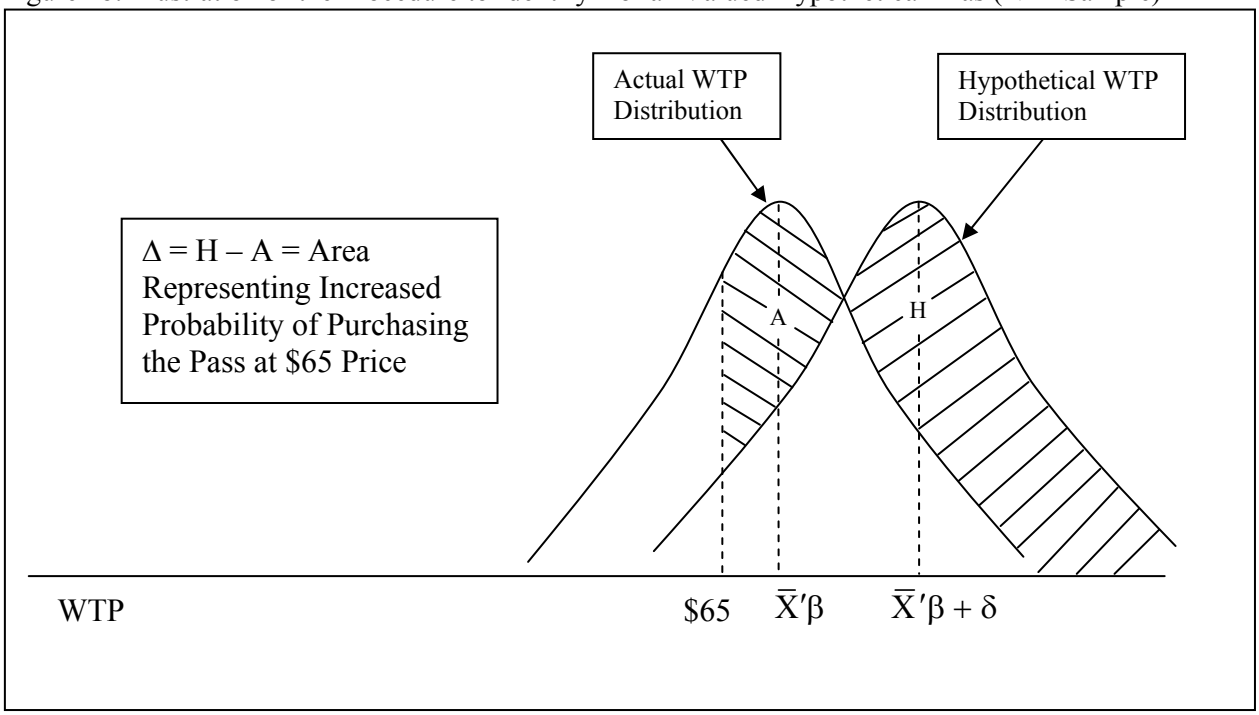


Figure 1b. Illustration of the Procedure to Identify Dollar-Valued Hypothetical Bias (NPF Sample)



Endnotes

[1] Although they are the exception rather than the rule, there are several empirical studies that have contrasted revealed and stated preference information for the same good or program (e.g., Adamowicz *et al*., 1994; McConnell *et al*., 1999; Cameron *et al*., 2002; Aadland and Caplan, 2003; Azevedo *et al*., 2003). Other papers have combined revealed and stated preference data with an experimental setting (e.g., Cummings *et al*., 1995; List, 2001) or investigated the theoretical implications of combining RP and SP data (e.g., Hensher *et al*., 1999).

[2] One minor difference is that the IAP card can be signed by two adult pass holders, either of whom may show the card to secure entrance, whereas the GEP card had space for only one signature. As a practical matter, however, this difference is inconsequential, because typically the individual showing the GEP on behalf of a carload of entrants was not required to verify that the signature was his or hers.

[3] We note that some may find an offsetting inconvenience in having to remember to bring the pass for each visit. In concept, some of the values associated with "convenience" also accrue to the federal government if having a pass reduces wait times at entrance stations, allowing fee collectors to work more efficiently.

[4] One phone number that was drawn in the RDD sample also appeared on the NPF population list. To maintain strict independence of the two strata, that household was dropped from the RDD sample; by chance, it happened not to be drawn in the NPF sample from the population list, but it could have been.

[5] Within-household sampling was not undertaken, because the relevant unit of analysis is the household. The IAP (like the GEP and NPP) is not an individual pass; rather, it entitles members of a household traveling together to access federal lands for recreation. Therefore, any willing adult respondent in each household included in the survey was asked to provide information on behalf of the household.

[6] While a larger final sample would have been desirable, the extensive screening just described was essential for analytic purposes. As will be evident in the results below, both of the screened sample sizes provide adequate statistical power to identify as "significant" the effects central to our analysis. The final response rates were good (65% for RDD and 55% for NPF, using the RR3 formula of the American Association for Public Opinion Research (2006)). Differences between weighted and unweighted descriptive results were small, suggesting that non-response bias was low. The weighting itself introduced little loss in statistical power (inflating standard errors only 7.5% in the RDD sample and less than 1% in the NPF); these power losses have been appropriately taken into account in our tests of statistical significance (using the method described by Dorofeev and Grant, 2007, p. 105).

[7] The linear specification in (1) allows for negative WTP values. This is consistent with the fact that some (non-protest) households are included that refused a free pass (without rejecting the survey premise) and would therefore need to be compensated to accept the pass. On the exclusion of protest households, see endnote 9.

[8] In other words, if the first bid price is accepted, the second price is (randomly) either $20 higher, or else (with equal probability) it is approximately double the first bid amount. If the first bid price is declined, the second price is either $20 lower, or else it is approximately half the starting amount. Adjustments of $5 (when doubling or halving the first bid) ensure that all bids are multiples of 5, to avoid any confounding that might arise if multiples of 10 have a different psychological resonance with respondents.

[9] A small number of protest households are excluded from all analyses, based on the original screening decision to omit households that are unlikely to participate in the market for the IAP. Protest households (N=6 for RDD sample, N=5 for NPF sample) are defined as those who not only refuse the pass for free, but also (when asked to explain) give open-ended answers rejecting the overall premise that federal government should be selling passes to public lands. Protest households made statements such as "I pay taxes; I shouldn't have to pay to get onto my federal land."

[10] We note that distance to federal recreation sites may not be entirely exogenous, if access to outdoor recreation is a factor in household decisions about location of residence. As a practical matter, measurement of a distance variable would be difficult for various reasons: (i) with hundreds of national park units (including urban historic sites) and thousands of other federal land sites, it is extremely difficult to identify the "nearby" sites relevant for each potential respondent; (ii) many federal recreation sites (especially national forests) have multiple points of entry, and it is not

uncommon for fee requirements to be enforced at some but not all entry points to a particular site; (iii) the definition of "nearby" is arbitrary; (iv) not all recreation sites are equally attractive recreation areas; etc.

[11] Exceptions exist—some experiments have found no significant difference in real and hypothetical behavior (e.g., Dickie *et al*., 1987; Loomis *et al*., 1996; Smith and Mansfield, 1998).

[12] Because our estimation of hypothetical bias captures *any* difference between stated and revealed preferences, it will include other differences such as strategic bias. Strategic bias occurs when households intentionally say 'NO' to an acceptable bid in order to encourage policymakers to set a lower price.

[13] We excluded 61 and 336 revealed-preference respondents in the RDD and NPF samples who were aware of the policy that allowed receipts from recent entrance fees to federal recreation sites to be applied toward the cost of the NPP or GEP. This was done to level the playing field because this unpublicized "receipt policy" may alter the value of a pass and was not described to survey respondents.

[14] Because we only use the first bid for the RP/SP comparison, we are not concerned with the issues of anchoring and structural shifts. The underlying WTP is therefore based on equation (1) rather than equation (4).

[15] We originally examined the identical set of control variables for the overall WTP model (1) and the hypothetical bias model (5). However, the race variables and most of the region variables were not significant predictors in (5) so we chose to simplify the hypothetical bias probit model presented here.

[16] An alternative procedure for identifying δ would be to multiply the estimate of σ from Table 4 times the estimate of (δ/σ) from Table 5. The drawback to this approach is that it only uses information from the conditional mean of WTP whereas our procedure uses information from the entire WTP distribution. The two methods give similar results for the NPF sample, but for the RDD sample this alternative would produce a dollar figure more than double what we find. The external validity check reported in Section 6 shows that our approach is only slightly conservative and yields an adjustment to the demand curve that corresponds quite well with subsequent pass sales.

[17] Personal communication (6/16/09) from Kim Magraw, Office of Policy Analysis, U.S. Department of the Interior.

[18] Figures on actual sales are ordinarily accumulated on a fiscal year basis; however, when the IAP first went on sale in January, 2007, three months of the federal fiscal year had already elapsed. To approximate an annual figure, we inflate by one-third the number of new passes sold in the first nine months of fiscal 2007. For comparison, we report sales covering the full twelve months of fiscal 2008. We also provide sales figures for both years on another pass, available only to senior citizens, although it was expressly excluded from our modeling. This pass changed its name, but not its $10 price, when the IAP went on sale. In Table 7, the figure for "Senior" passes sold during fiscal 2007 includes three months of sales under its former name.

[19] The demand curve is derived by counting the number of sample households with a predicted WTP higher than a given price and then scaling the resulting figure to the population of potential IAP purchasers.