ISSN 1047-3289 *J. Air & Waste Manage. Assoc.* **55**:1767–1779

Methods and Results from a New Survey of Values for Eastern Regional Haze Improvements

Anne E. Smith, Michael A. Kemp, Timothy H. Savage, and Catherine L. Taylor
CRA International, Washington, DC

ABSTRACT

A contingent valuation survey was designed and implemented in 2003 to elicit the willingness to pay for haze reduction in scenic areas of the Eastern United States. This survey updated and enhanced a portion of a 1988 survey, Chestnut and Rowe's Preservation Values Study, which is commonly used to estimate benefits of haze reduction policies. The 2003 survey incorporated features to address criticisms of the 1988 survey regarding its lack of budget constraints and its simplistic descriptions of changes in visibility. In preparing the new survey, we identified a fundamental error in the 1988 survey regarding the amount of visibility change that respondents were asked to value. When corrected, it reduces the estimates in that study of willingness to pay per unit of visibility change by ~40%. More importantly, we estimate that haze reduction benefits based on that study are overstated by ~70%. Results from our 2003 survey demonstrate that the distribution of willingness to pay is highly skewed, which implies that a haze reduction policy of which the costs per person are equal to the average willingness to pay would be rejected by between 70% and 80% of the public if subjected to a majority vote. Moreover, estimates of willingness to pay are sensitive to how respondents are reminded of their budget constraints. We find a statistically significant decline in willingness to pay when budgetary reminders are included. Finally, we find evidence that a large portion of the willingness to pay appears to be associated with improvement on the worst condition days. Little additional value appears to be associated with improvements to the majority of days that have haze conditions nearer the average. These survey results have important implications for policy making on haze reduction.

IMPLICATIONS

The Regional Haze Rule (40 *Code of Federal Regulations* 51.308) requires policymakers to choose rates of visibility improvement for each implementation period, which requires balancing benefits of improved visibility against costs. This paper presents initial results from a new contingent valuation survey on U.S. willingness to pay for visibility improvements that can inform policy analyses better than previous valuation surveys. This study highlights previous and extant biases in visibility valuation estimates and provides other insights on complexities and nuances in inferring societal willingness to pay estimates that also should be incorporated into policy deliberation.

INTRODUCTION

Stringent new air quality regulations, such as the Regional Haze Rule (40 *Code of Federal Regulations* [CFR] 51.308), have created a need for reliable and policy-relevant estimates of visibility benefits. To that end, a new contingent valuation (CV) survey was conducted in 2003 to estimate willingness to pay (WTP) for changes in visibility conditions in Eastern U.S. national parks. Economists typically use CV as a tool to value nonmarket goods, where WTP is elicited directly through a survey. More comprehensive information on the CV method can be found in Mitchell and Carson[1] and Bateman and Willis.[2]

Our research took as its starting point an earlier CV study that has been cited widely in the analysis of regional haze, "Preservation Values for Visibility Protection at the National Parks" by Chestnut and Rowe.[3] The paper by Chestnut and Rowe has been used to estimate visibility benefits associated with air quality programs, including the fine particulate matter (PM) and ozone National Ambient Air Quality Standards, the Regional Haze Rule, and the report to the U.S. Congress on the prospective costs and benefits of the entire Clean Air Act.[4–6] In this survey, we sought to replicate the Chestnut and Rowe study and to improve on its methodology by incorporating new insights of CV research since 1988, the year it was conducted.

A major task in the development of the 2003 survey was to prepare a set of photographs and histograms to define the visibility commodity to be valued. Scenarios of visibility change were constructed to emulate the scenarios of change used in the paper by Chestnut and Rowe. This survey, however, provided respondents with more thorough descriptions of the nature of visibility change. In particular, it provided detailed information on the full distribution of daily visibility conditions and how those conditions change in each scenario. In preparing this survey, it became apparent that the original Chestnut and Rowe photographs, used to display changes in visibility, did not accurately reflect their intended distributions. As we describe here, the technical error in Chestnut and Rowe's representation of baseline conditions has created a substantial upward bias in estimates of visibility benefits.

We describe the types of questions that were asked in this survey, as well as estimates of WTP. In addition to replicating Chestnut and Rowe's survey, this survey explored whether various reminders of budget constraints, as in Kemp and Maxwell,[7] affect a respondent's WTP. One version of our questionnaire simply encouraged respondents to consider both their personal financial circumstances and other public policy priorities before articulating a WTP value. Another questionnaire affirmatively presented competing alternatives to air quality improvements when assessing WTP for haze reduction, thereby disaggregating competing budget demands.

Defining the Visibility Scenarios Valued in the Survey

In 1993, the *Federal Register* published a report on good practices for producing reliable WTP estimates from CV, known as the "Blue Ribbon Panel Report."[8] One issue cited was the need to provide sufficient information in the survey to ensure that respondents understand as precisely as possible the commodity they are asked to value. Visibility is a notoriously difficult environmental commodity to define for a CV survey. The public typically has no experience assigning quantitative measures to various degrees of visibility impairment, because the metrics that scientists use to characterize haze (light extinction, visual range, and deciviews) are largely meaningless to the general public. Lacking a meaningful way to describe visibility changes, CV surveys of visibility must rely on photographs. While addressing the difficulty of defining visibility change, photographs present other difficulties because: (1) they may narrow the focus to certain vistas at particular points in time; and (2) they may introduce confounding variations in other esthetic attributes because of variations in scenery at the time the picture is taken (temporal variability) and to variations in coloration and focus that arise during the production process. The latter problem can be mitigated by using computer-based methods to generate precise amounts of haze against a common base photograph and by conducting careful quality control in preparing the visuals used in the survey. Computer-generated photographs were a feature of the 2003 survey.

However, even computer imaging does not solve the complex problem of temporal variability. Visibility, a purely esthetic amenity, is experienced as a continuous sequence of near-instantaneous events. A continuous, real-time depiction of visibility, through use of video, for example, is not practical in a large-scale survey. Chestnut and Rowe opted to inform survey respondents on the then-current visibility conditions using a set of four still photographs that depicted the same vista during the summer. Each of the four photographs reflected a perceptibly different amount of haze. For eastern conditions, Chestnut and Rowe used a vista from the Shenandoah National Park in Virginia, called "Rocky Mt." We adopted the same vista to limit differences between the new results and the Chestnut and Rowe results to well-controlled differences in the survey design.

The haze depicted in each of Chestnut and Rowe's photographs was intended to represent a specific point along the then-current distribution of daily summertime visibility. Chestnut and Rowe, however, limited the characterization of the distribution to descriptions of the current frequency for each of the four photographs. Specifically, beneath each photograph (labeled A through D, with A being the clearest and D the haziest) were captions indicating that they occurred on 15%, 20%, 40%, and 25% of summertime days, respectively. With these particular frequencies, photograph C represented mean conditions, and Chestnut and Rowe told respondents that photograph C was typical of average conditions. Thereafter, the descriptions of the visibility commodity in the paper by Chestnut and Rowe were stated in terms of average conditions. For example, respondents were asked to value policy changes that would improve average visibility to conditions represented in photograph B. In turn, they were also asked their WTP to improve average visibility to that of photograph A and to avoid deterioration to average visibility like that of photograph D.

With the use of these photographs, the Chestnut and Rowe survey weakly characterized visibility as the cumulative effect of a variety of daily experiences. In fact, their characterization of visibility changes implicitly encouraging respondents to assign WTP values based on two-way comparisons of photographs, that is, C versus B, C versus A, and C versus D. Chestnut and Rowe did not specify changes in overall distributions, and one must interpret their survey results without knowing what respondents had assumed about day-to-day changes under each scenario. We believe that a reliable CV-based estimate of WTP requires a more detailed and unambiguous description of how respondents should expect their daily experiences of visibility to be affected. This survey improved the characterization of policy impacts by adding more explicit information on the overall distributions and by specifically describing each scenario of change as a shift from the baseline distribution. It did not tell respondents the average in each scenario and, thus, could not prompt respondents to make simplistic comparisons of just two photographs when considering the differences they were meant to evaluate.

We now explain how we sought to achieve the goal of developing our survey by the following: (1) constructing a baseline visibility distribution from first principles; (2) characterizing each policy scenario as a full distributional shift; and (3) condensing complex distributional information into an integrated set of photographs, graphical displays, and text that respondents could easily assimilate.

Derivation of Baseline Visibility Distribution

A careful characterization of the visibility distribution from first principles required an analysis of a large record of air quality and weather data. We used data from the Interagency Monitoring of Protected Visual Environments (IMPROVE) monitoring network maintained by the National Park Service.[9] This monitoring network maintains a continuous record of ambient aerosol concentrations, weather conditions affecting light extinction, and directly measured light extinction in many U.S. national parks. The two direct measures of light extinction are drawn from nephelometers and transmissometers. However, nephelometers measure light scattering by particles but not light absorption, and IMPROVE strongly cautions against using only transmissometer data as a measure of visibility.[10] Therefore, we used reconstructed light extinction for our analysis. Given our use of the Shenandoah vista used by Chestnut and Rowe, we relied on data from the Shenandoah National Park monitor, which extends back to 1988. We also reviewed data for other eastern monitor locations in the IMPROVE network and found that the Shenandoah data are representative of eastern conditions generally.

Many different types of ambient aerosols and atmospheric gases cause light extinction, each with a different degree of extinction per unit of ambient concentration. Additionally, light extinction is enhanced by relative humidity (RH), which makes some particles larger. Sulfates (SO_4) and nitrates (NO_3) are the sources of extinction most strongly affected by RH. Although there is uncertainty in the specific parameters, researchers currently use the following equation from IMPROVE[11] to "reconstruct" light extinction from information on ambient concentrations, measured in micrograms per cubic meter, of key aerosols:

$$\begin{aligned}\text{Light Extinction} = {} & 10 + 3*(\text{concentration of ammonium sulfate + nitrate})*F_T(RH) \\ & + 4*(\text{concentration of organic mass}) + 1*(\text{concentration of soil}) + 0.6*(\text{concentration of coarse matter}) \\ & + 10*(\text{concentration of light-absorbing carbon}) \qquad (1)\end{aligned}$$

The constant term, 10, is the light extinction associated with natural atmospheric gases, often called Rayleigh extinction. $F_T(RH)$ is a nonlinear function of the RH reflecting the fact that the extinction efficiency of SO_4 and NO_3 increases as RH increases, which is particularly steep for RH >80%:

$$F_T(RH) = -0.18614 + 0.99211*[100/(100 - RH)] \qquad (2)$$

We used the specific form of $F_T(RH)$ for summer months provided in Table 1 of IMPROVE.[10] It was developed for use with daily average RH values and used in conjunction with daily average aerosol concentrations as reported in the IMPROVE data.

Following Chestnut and Rowe, this survey was restricted to WTP for changes in summertime visibility. We developed the current summertime distribution by extracting from the IMPROVE aerosol concentration files[12] all of the summer days, July 1 through September 30, with complete data for each of the aerosols in the light extinction equation. For current baseline conditions, we aggregated data for 1997–2001, the most recent set of 5 yr for which complete, quality-assured summertime data were available. To understand how the baseline distribution has changed since the time of Chestnut and Rowe, as well as to verify Chestnut and Rowe's reported baseline distribution, we applied the same methods to the Shenandoah data for 1988–1992. 1988 is the earliest year for which aerosol data are available for the Shenandoah site, and Chestnut and Rowe's survey was conducted in 1988. We developed our estimate of each distribution by treating each observation extracted in the above manner as an independent random draw from an underlying summertime distribution. We treated all of the observations as individual points along this underlying distribution rather than fitting a parametric distribution.

Estimation of light extinction also required data on RH. For regulatory purposes, U.S. Environmental Protection Agency (EPA) mandates use of a specific, fixed set of average monthly RH values to produce a set of daily reconstructed light extinction values, which IMPROVE reports with each of its daily aerosol concentration observations. However, RH varies widely across the days within a month, and the method of setting RH at a fixed value reduces the variance from the observable distribution of light extinction. The baseline distribution in a WTP survey should reflect all of the important sources of variation and not just a regulatory definition of haze conditions. For this reason, we combined the aerosol concentrations of IMPROVE for each day with a distribution of RH values. Data supporting our distribution of summertime daily average RH values were also taken from the IMPROVE data for the Shenandoah site. With these two sources of variance, we then created our own distribution of daily average light extinction, using the same equation that EPA mandates for reconstructed light extinction, but also incorporating variance in daily average RH. The distributions for visual range and deciviews were derived from the light extinction distribution.

Crucially, our method preserved the correlation in actual concentrations of different types of air pollution caused by the fact that most species of air pollution tend to move together in air masses. Although it assumed RH and aerosol concentrations are probabilistically independent, our method uses a more realistic degree of variance in the baseline distribution of visibility than if we had relied on the reconstructed light extinction exactly as reported in the IMPROVE database. The tenth and ninetieth percentiles in our estimated distribution of non-rainy, summer daytime average RH were 60% and 80%, respectively. In contrast, the EPA formula for reconstructed light extinction assumes that RH is a fixed value in each month with a low of ~72% in June and a high of 76% in August. Given the nonlinearity of the F_T(RH) function, our method introduced a variation of approximately ±40% in SO_4 and NO_3 light extinction when compared with reconstructed light-extinction data in IMPROVE.

Figure 1 shows the resulting 1997–2001 distribution of visual range and deciviews for summer months, non-summer months, and all months. It highlights the degree to which the Shenandoah area experiences more visibility degradation during summertime than during the rest of the year. Average visibility in the summertime is 31 km or 25 deciviews (dv), whereas all of the other months have an average of 56 km or 20 dv. We compared the 1997–2001 summertime distribution to a comparably constructed distribution for the earlier period, 1988–1992, which would reflect the baseline distribution at the time of the Chestnut and Rowe survey. Whereas this revealed improvement since the mid-1980s as a result of improved air quality, the baseline distribution we used matches fairly well the distribution reported by Chestnut and Rowe.

Distributional Shifts to Emulate Chestnut and Rowe Scenarios

As noted above, Chestnut and Rowe's scenarios were not described in terms of how the baseline distribution of visibility would shift, but rather were described as changes in average summertime visibility. One version of the 2003 survey exactly replicated this method of presenting improvements in visibility. Another version presented changes in visibility as changes from the baseline distribution to new distributions. The new distributions had average changes comparable to those of the Chestnut and Rowe scenarios, but these variants did not tell respondents what the average was or how it was being changed.

Because there are multiple probability distributions that yield the same average, we applied a simple assumption that all of the days would experience an equivalent proportional increase in visual range (decrease in deciviews). To determine the proportional scalings for each scenario, we identified the average values of the baseline distribution for this paper within each of the probability ranges that Chestnut and Rowe used for their four photographs. Table 1 compares the values from the baseline distribution of this paper to those of Chestnut and Rowe. For Chestnut and Rowe photograph A, for example, the clearest 15% of days, the average visibility of the 2003

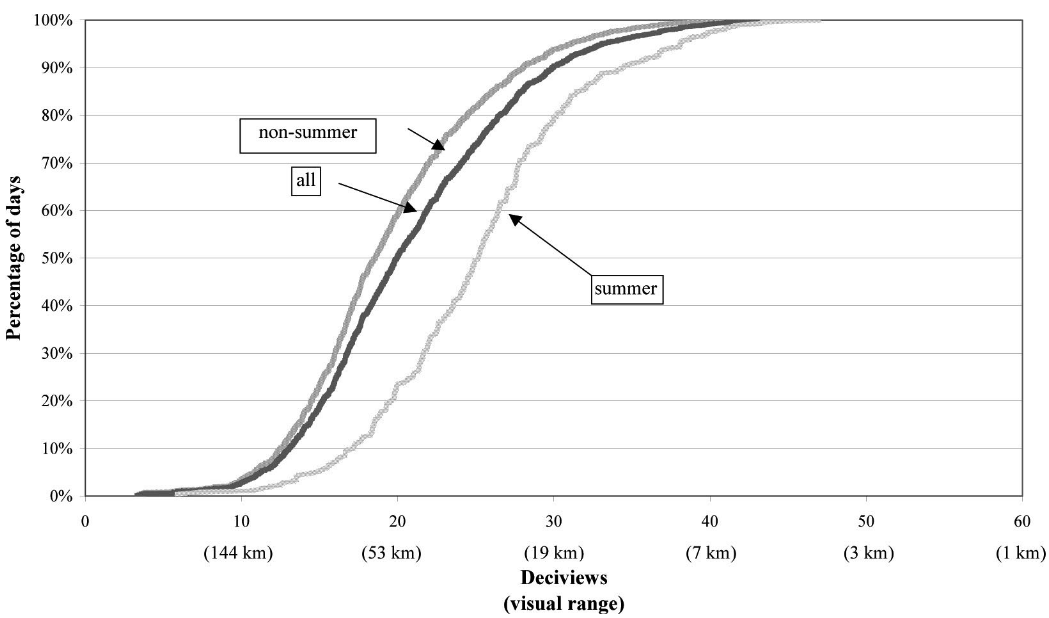

Figure 1. Current baseline distribution of deciviews derived from 1997–2001 IMPROVE data.

Table 1. Comparison of baseline visibility distribution described by Chestnut and Rowe to current visibility distribution estimated from IMPROVE data.

Photograph	Description	Average Visibility in Chestnut and Rowe Baseline	Average Visibility in Current Baseline
A	Clearest 15% of days (0–15th percentile)	75km (17dv)	85km (15dv)
B	Next clearest 20% of days (15–35th percentile)	50km (21dv)	50km (21dv)
C	Next clearest 40% of days (35–75th percentile)	25km (28dv)	30km (26dv)
D	Worst 25% of days (75–100th percentile)	10km (37dv)	13km (34dv)

survey is 85 km compared with an average of 75 km in Chestnut and Rowe paper. The next clearest 20% of days, which Chestnut and Rowe represented in photograph B, was 50 km on both baseline distributions. Table 2 summarizes the proportional scalings used to shift the baseline distribution of this paper by amounts that were consistent with each of the three average-change scenarios used by Chestnut and Rowe. Figure 2 displays the resulting distributions used in the 2003 survey, where β is comparable with the Chestnut and Rowe scenario in which the average has shifted to photograph B, γ is comparable with the Chestnut and Rowe scenario in which the average has shifted to photograph A, and δ is comparable with the Chestnut and Rowe scenario in which the average has shifted to photograph D.

Figure 2 depicts a fourth scenario, α, which was also included in the 2003 survey. This fourth scenario was of a type not considered by Chestnut and Rowe and entails a nonproportional shift from the baseline distribution. In this new scenario, only the worst visibility days are improved. To construct this scenario, the days in the top twentieth percentile of the distribution were rolled back by an amount that was randomly selected from the range between 0 and 15 dv. Therefore, this scenario represented a very large potential improvement on a relatively small portion of days. For example, the eightieth percentile day had a 30-dv degradation in visibility, and a 15-dv improvement on such a day implied a 350% increase in visual range. On the other hand, these large improvements did not occur every day. Summertime average visibility was reduced by only 1.4 dv, whereas the annual average is reduced by <1 dv. Chestnut and Rowe's method could not represent such a small change. However, in a survey that depicts visual change as distributional shifts, it was possible. Most importantly, this scenario was much closer in design to the literal requirements of the Regional Haze Rule.

Table 2. Characterization of distributional shifts commensurate with three visibility change scenarios used by Chestnut and Rowe.

Average Change (scenario in 2003 survey)	Proportional Rollback Factor Applied to New Baseline Distribution	Change in Visual Range (change in dv)
C to B (β)	0.61	37% increase (−5 dv)
C to A (γ)	0.82	135% increase (−10dv)
C to D (δ)	1.35	64% reduction (+9 dv)

Photographs and Histograms Used to Communicate the Scenarios in the Survey

Photographs rather than words were the primary method of relaying the various visibility conditions. Neither Chestnut and Rowe nor this survey mentioned visual ranges, deciviews, or light extinction to respondents. Rather, four photographs were provided to illustrate visibility at specific points on the distribution. However, several changes were made in this survey to improve the visuographic characterization of the visibility commodity.

The authors obtained the original slide of Chestnut and Rowe's photograph A, the clearest day, and had it converted to a digital format. All of the photographs presented in this survey were created from this base photograph using computer-imaging software[13] to precisely reflect four specific levels of visibility along the distribution as uniform changes in light extinction. A graphic like the one shown in Figure 3 was provided to survey respondents along with the photographs to characterize the change in visibility conditions that they were being asked to value. This graphic includes two histograms showing the frequency and number of summertime days that would be associated with conditions like those in each photograph. Each scenario was characterized by showing

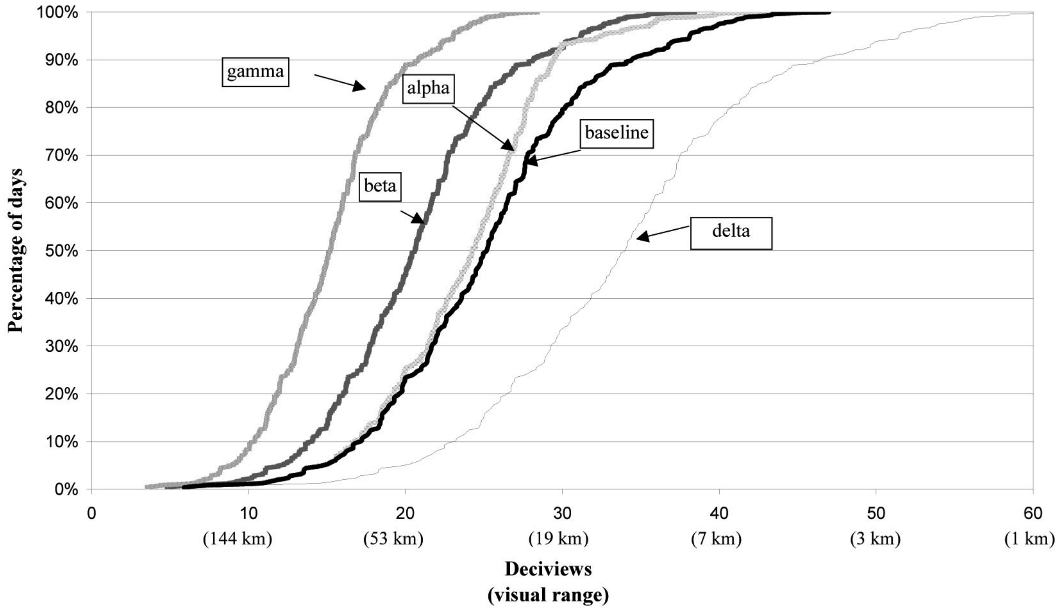

Figure 2. Full distributions for baseline and four scenarios of changed visibility conditions in 2003 survey.

the baseline histogram beside the histogram of the new scenario. Both of these were placed beneath the photographs with each bar of the histograms arrayed in the same horizontal ordering as the photographs. A fifth "blank" photograph was added to the sequence to reflect the fact that some days would be rainy or foggy when visibility differences caused by air pollutants would be imperceptible. The histograms included a fifth bar to reflect the frequency of these weather-impacted days.

Freed by the histograms from the need to have one photograph represent "average" conditions, the four photographs of this survey reflected visual ranges more evenly spaced along the probability distribution. Scenarios were no longer associated with particular photographs; instead, the histograms revealed the scenarios as changes in the frequencies of days represented by each of the four fixed photographs. Table 3 summarizes the visual range for each photograph and their respective frequencies under the baseline and each of the scenarios.

Errors in Chestnut and Rowe Photographs and Resulting Biases in Visibility Benefits Estimates

The decision to use computer-imaging techniques to prepare the photographs in this survey, while still relying on a base photograph in the Chestnut and Rowe survey, revealed an error in that survey. Chestnut and Rowe states that their four photographs were supposed to depict the Shenandoah vista under visual ranges of 75, 50, 25, and 10 km, respectively. However, photographs generated by the imaging software for these visual ranges did not match the apparent haziness in the Chestnut and Rowe photographs. Therefore, we used the imaging software to identify the best possible match to the haziness in the Chestnut and Rowe photographs and determined that they were much closer to 150, 50, 17, and 5 km, respectively. Air Resource Specialists, Inc., which provided the imaging software, confirmed that Chestnut and Rowe had assigned incorrect visual ranges to their photographs by using additional evidence regarding specific elements in

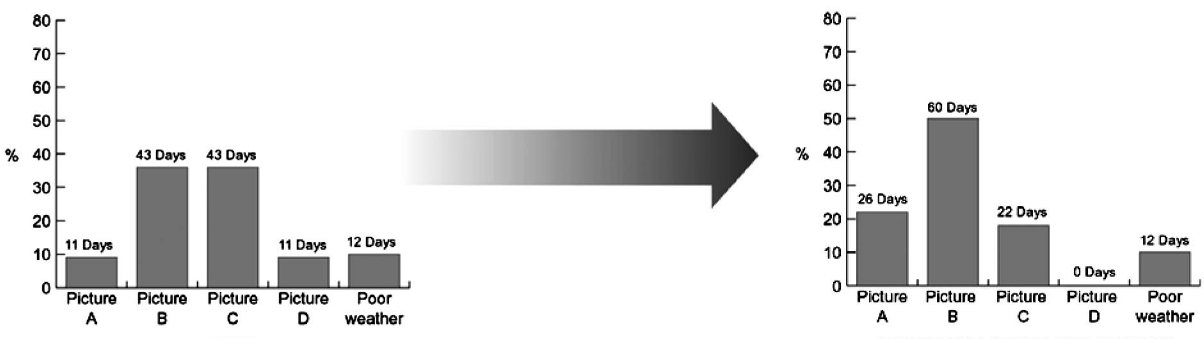

Figure 3. Example histogram used to illustrate the β scenario.

Table 3. Summary of baseline and four alternative visibility distributions in terms of frequency for each photograph used in new survey.

Variable	Photo A 96 km 14 dv (%)	Photo B 45 km 22 dv (%)	Photo C 23 km 28 dv (%)	Photo D 8 km 38 dv (%)	Blank[a] (%)
Baseline	9	36	36	9	10
α scenario	9	39	40	2	10
β scenario ('C to B')	22	50	18	0	10
γ scenario ('C to A')	72	18	0	0	10
δ scenario ('C to D')	3	5	37	45	10

[a]Reflecting too much rain or fog to discern air pollution impacts to visibility.

the photographs and known distances to them from the viewing point. Figure 4 illustrates the substantial difference between the distribution Chestnut and Rowe attempted to reflect and the distribution that they actually presented to respondents with their photographs. It was the latter distribution on which Chestnut and Rowe's respondents based their WTP responses. Their distribution actually valued a much larger variance than the baseline Chestnut and Rowe reported it had used. Thus, the changes in average visual range to which the WTP responses apply were much larger than Chestnut and Rowe reported.

For the version of our survey that exactly replicated Chestnut and Rowe's survey, we used photographs with the 150, 50, 17, and 5 km visual ranges to closely match the images Chestnut and Rowe actually used. Because the Chestnut and Rowe questionnaire never mentioned any numeric metrics to describe the visual ranges in the photographs, the closest replication would be to show the same set of visual ranges. The Chestnut and Rowe error, however, had two implications for the interpretation of their WTP estimates. First, the Chestnut and Rowe survey described baseline conditions inaccurately, overstating the variance of the then-current distribution. This might have induced respondents to perceive a greater WTP for a given amount of change in visibility. Second and more importantly, it created an unambiguous upward bias in interpretations of the Chestnut and Rowe results as WTP per unit of visibility change. This created a large overstatement in benefit estimates based on their results used to support policy proposals.

It is possible to quantify the magnitude of the latter overstatement, because it only affects ex-post interpretations of the WTP results. For example, changing average visibility from that depicted in photograph C to that in photograph B was reported in Chestnut and Rowe as a 25-km change (a 100% increase in visual range), when it was actually a 33-km change (a 200% increase in visual range). Table 4 reports summertime average visibility change before and after the error is corrected. Simple inspection of Table 4 indicates that benefit estimates based strictly on the data reported in Chestnut and Rowe have probably been overstated by 50–100%, whether assessed as percentage changes in visual range (which is the way Chestnut and Rowe's results have been interpreted in the benefits analysis literature) or as numbers of deciviews changed (a more recent way of considering the "amount" of change in visibility conditions). This overstatement

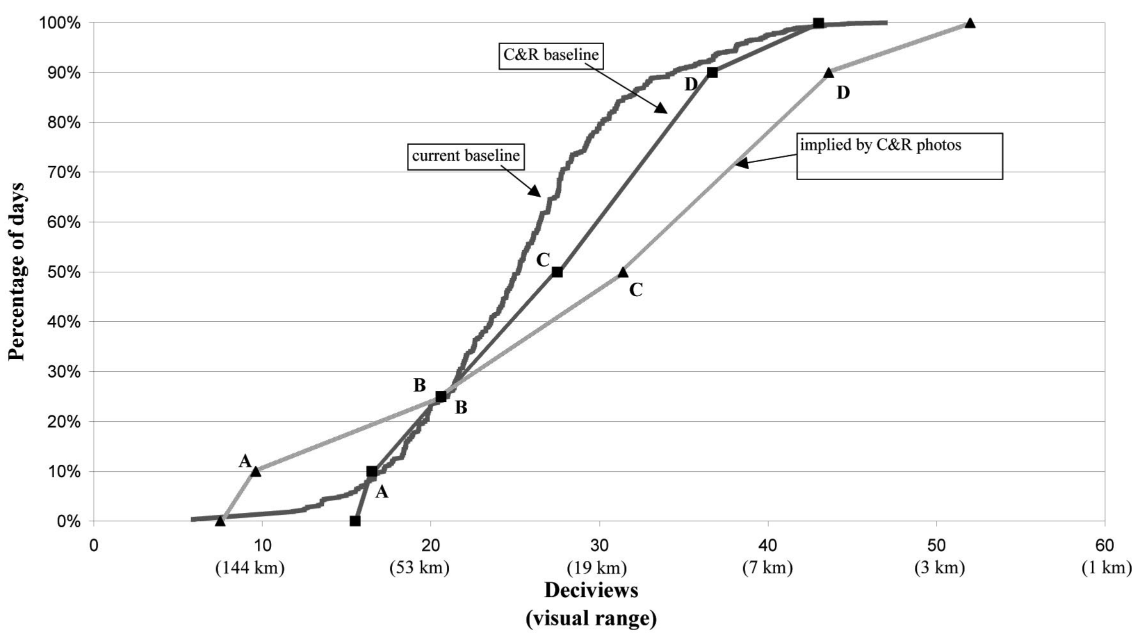

Figure 4. Comparison of current base conditions with the reported Chestnut and Rowe baseline and the baseline implied by the Chestnut and Rowe photos.

Table 4. Changes in average visibility reported by Chestnut and Rowe for each of their WTP estimates for Southeast compared with average changes actually depicted to survey respondents.

Average Change (scenario in 2003 survey)	Changes in Visibility Reported Reported by Chestnut and Rowe	Changes in Visibility Reported Used in Chestnut and Rowe Survey
C to B (β)	+100% (−7 dv)	+190% (−11dv)
C to A (γ)	+200% (−11dv)	+780% (−22dv)
C to D (δ)	−60% (+9 dv)	−70% (+12dv)

would filter through to numeric comparisons in meta-analyses, such as Smith and Osborne,[14] which include Chestnut and Rowe estimates.

The degree of overstatement can be quantified even more precisely by turning to Chestnut and Dennis.[15] This paper used the Shenandoah vista WTP results from Chestnut and Rowe to extrapolate to benefits in the Eastern United States attributable to sulfur dioxide reductions under Title IV of the Clean Air Act Amendments of 1990. It introduced a method of extrapolation where a benefits function was fit to the WTP results for the three scenarios of average visibility changes valued in Chestnut and Rowe:

$$WTP = B*\ln(VR2/VR1) \tag{3}$$

where VR2 is the average visual range after the change, and VR1 is the average visual range in the baseline. The term VR2/VR1 is equal to the percentage change in Table 4 plus 100%. Using the percentage changes reported in Chestnut and Rowe (i.e., those in column 2 of Table 4), Chestnut and Dennis estimated a value of B = 50 for non-Virginians and B = 85 for Virginians. However, when we re-estimated B using the percentage changes associated with the actual photographs used in the paper by Chestnut and Rowe (i.e., those in column 3 of Table 4), we obtained a value of B = 30 for non-Virginians and B = 50 for Virginians. Thus, the benefits that Chestnut and Dennis estimated for visibility improvements in national parks of the East under the Title IV program, based on the original Chestnut and Rowe WTP values, were overstated by ~70%.

The Chestnut and Dennis method was also used in major visibility benefits studies by EPA.[4–6] More recently, EPA shifted to a more complex benefits function based on a constant elasticity of substitution utility function;[5,6] however, the parameters of this benefit function also have been calibrated to the Chestnut and Rowe WTP results. Thus, all of these other important benefits analyses are also affected by the visual range errors in Chestnut and Rowe, and their estimated dollar benefits are likely overstated by a similar magnitude.

Methodology of New Survey

One goal of this study was to replicate the Chestnut and Rowe survey as closely as possible to assess changes in the public's valuation of visibility improvements that have occurred since its implementation, independent of any methodological improvements. Another goal was to incorporate new insights and recent findings of CV research to produce a state-of-the art CV survey and WTP estimates. Additionally, the authors wished to value a wider variety of visibility changes than Chestnut and Rowe. To achieve these goals we fielded multiple versions of the survey, described below.

Version 1 replicated closely that part of the Chestnut and Rowe study concerned with visibility in Eastern U.S. national parks. The only major difference was that it used computer-imaged versions of the Chestnut and Rowe photographs. The photographs in version 1 reflected the same visibility conditions actually shown in Chestnut and Rowe, rather than the visibility conditions Chestnut and Rowe reported numerically. We had no expectation that this version would necessarily generate mean WTP values similar to Chestnut and Rowe for at least two reasons. First, valuations among the population have changed in the intervening 15 yr. Second, the sample design and mode of survey administration in this paper were significantly different than Chestnut and Rowe. Chestnut and Rowe used a self-completion questionnaire that was mailed to residents of only five states. The authors of this paper used a computer-assisted online questionnaire that was distributed to a randomly selected national panel.

Version 2 is identical to the version 1, except for the following reasons: (1) the photographs were changed to correctly represent current baseline visibility conditions; (2) "bad weather days" were represented explicitly; and (3) changes in visibility in each photograph were presented as shifts in distributions rather than as changes in the overall "averages."

Version 3 is identical to version 2, except that it loosened the close ties to the questionnaire language used by Chestnut and Rowe. Specifically, the authors lengthened the description of the scenario, incorporated new questions, and augmented wording intended to strengthen respondents' awareness of budget constraints when responding to WTP questions.

Version 4 is identical to version 3, except that it adds a scenario, α, which characterizes a reduction in haze on the worst days but with little change in average conditions. In the empirical analysis, the results from versions 3 and 4 are combined.

Version 5 provides a more radical deviation from the previous versions. Respondents were asked to express a total WTP for air quality improvements generally. They were then asked to allocate that total WTP across three different programs: one targeted on the reduction of deleterious health impacts; one focusing on reducing harm to plant life; and the third addressing haze effects on the visibility of scenic vistas. The latter program was characterized by two of the change scenarios used in version 3.

Finally, the 2003 survey, including version 1, contained a more extensive battery of debriefing questions than did Chestnut and Rowe. Copies of the 2003 survey and all of the survey materials are available from the authors on request.

Survey Fieldwork

The sample for this survey was drawn from the randomly selected, prerecruited national panel maintained by Knowledge Networks, Inc. Unlike other established panels, Knowledge Networks does not accept volunteers, but instead uses random-digit dial telephone survey methods to screen and recruit participants to the panel. The company's surveys are designed for online administration. It provided free WebTV units to recruited panel members or allows them to use their own internet-accessible equipment to complete a survey for which they have been selected. Information on the units is available from http://www.msntv.com. The advantages of using a prerecruited panel were primarily: (1) the ability to match total population demographic and socioeconomic characteristics very closely in the achieved sample by virtue of the availability of very large numbers of prescreened respondents and substantial experience in anticipating response rates for respondents of different types; and (2) a relatively high-response rate, as measured by the ratio of the achieved to the issued sample for any specific survey of panel members, thereby reducing the risk of nonresponse biases.

The sample of residents of the 48 contiguous states was designed to reflect key Census household population distributions by age, sex, race/ethnicity, education levels, and Census region. After substantial development work—three pilot rounds of in-person or online interviews combined with qualitative debriefing discussions—the survey was fielded to 2770 panel members commencing on October 30, 2003, and ended 3 weeks later. Approximately 1 week before the survey was fielded, each sampled panel member received a U.S. priority mail envelope containing an introductory letter alerting him or her to the imminent survey and its potential importance to policymakers. For sample members selected to receive versions 1 or 2 of the questionnaire, the letter followed Chestnut and Rowe in explicitly mentioning national parks as the focus, whereas the letter to the samples for the remaining versions spoke more generally of a survey about "national priorities." The new visibility survey was administered primarily as an online interview, except that the visual aids were mailed in advance in hard copy to ensure that all of the respondents used controlled, identical visual stimuli. Photographs were not displayed on the respondents' computer monitors or WebTV screens because they would have been of variable resolution, coloration, and contrast, thereby introducing uncontrollable sources of variation.

The survey achieved a 73% completion rate by the closeout date. After eliminating respondents who did not fit the sole screening criterion as a head of household, the authors achieved 1863 completed surveys. The authors defined head of household as "one of the people who makes most of the decisions about major purchases and other financial matters." The achieved sample was then weighted independently by questionnaire version to more closely mirror the sample universe by education, age, sex, and region of residence. Note that the authors' sample design, like their method of survey administration, differed markedly from Chestnut and Rowe.

Editing WTP Responses

After presenting the various scenarios, the authors asked respondents to choose a dollar amount from a provided list that represents the amount they would be willing to pay to gain the specified improvement (or avoid the specified degradation). This method is called the "payment card elicitation" method. The completed response records were edited in ways that are customary in CV surveys. The primary focus was on the WTP questions. As always, some respondents were unable or unwilling to answer these questions. Chestnut and Rowe report that ~7–8% of their respondents omitted WTP responses. The comparable statistics in this study were 15% omissions for versions 1 and 2 (Chestnut and Rowe replications) and 7% for versions 3, 4, and 5. The difference among the versions might be because the innovations of this study in the latter versions provide more information on the nature of the visibility changes. Also, the online questionnaire of this study provided an explicit "I don't know" option for WTP responses, whereas the Chestnut and Rowe questionnaire did not. It was, of course, possible for Chestnut and Rowe respondents to write "I don't know" or to skip a question.

The authors scrutinized zero WTP responses to identify those judged to be "protest zeroes." The respondents may have chosen an explicit $0 response for reasons other than that they place no value on the environmental good. They might not have fully understood the question they were being asked; they might have objected to details of the described payment scenario (cost-of-living increases deriving from higher prices and taxes); they might not have believed that the described change in visibility was feasible; or they might have doubted that governments would manage the program effectively. As in Chestnut and Rowe, zero values given for any of the above reasons were designated as protest zeroes and recoded as missing for the summary measures displayed in Table 5. Chestnut and Rowe retained nearly 70% of the zero bid amounts, whereas this paper keeps ~72%.

The authors also followed Chestnut and Rowe in eliminating seemingly abnormally high WTP amounts that might be the result of strategic bidding. Chestnut and Rowe flagged 3% of their respondents who either gave an adjusted WTP value >1% of their reported annual pretax income or selected the highest precoded WTP value for all three of the WTP questions. However, after examining the responses to an open-ended question posed to the high-end respondents, Chestnut and Rowe eliminated only four of the flagged responses. On the other hand, the authors of this paper examined two sensitivity analyses for measures of WTP. The first method, which the authors characterize as an "income trim," eliminated all of the responses >1% of the reported pretax income, reducing the sample sizes of key subgroups by 4–8%. The second method, a 5% trim, removed the highest 5% of WTP responses, which the authors report in Table 5. The multivariate regression models that we used to examine "conditional" means directly control for income and, therefore, do not involve any such trimming of data.

At the time of its publication, the Chestnut and Rowe study was innovative in that it included debriefing questions intended to gauge explicitly the extent to which respondents may have overestimated their WTP values. It

Table 5. Estimates of willingness to pay ($, 2003).

Variables	Chestnut and Rowe	Version 1 (Chestnut and Rowe replication)	Version 2 (add histograms)	Versions 3 and 4 (add budget context)	Version 5 (disaggregation)
α scenario					
Mean deciview change				1 dv	
Mean				$42±$13	
Median				$4	
Mean (5% trim)				$21±$5	
Respondents (*n*)				394	
Mean of predicted values				$16±$10	
Pooled respondents (*n*)				401	
β scenario					
Mean deciview change	11 dv	11 dv	5 dv	5 dv	5 dv
Mean	$48	$77±$22	$69±41	$59±$17	$24±$7
Median	$19	$35	$10	$10	$7
Mean (5% trim)		$55±$14	$35±$8	$33±$7	$12 ± 3
Respondents (*n*)	226	217	225	387	186
Mean of predicted values		$62±$25	$54±$40	$33±$15	$9±$8
Pooled respondents (*n*)		1017	1017	1017	1017
γ *scenario*					
Mean deciview change	22 dv	22 dv	10 dv	10 dv	10 dv
Mean	$75	$118±$55	$106±$67	$76±$15	$30±$10
Median	$25	$50	$20	$13	$9
Mean (5% trim)		$70±$20	$52±$14	$39±$6	$18±$5
Respondents (*n*)	226	222	231	784	177
Mean of predicted values		$101±$56	$84±$66	$34±$13	$14±$10
Pooled respondents (*n*)		1418	1418	1418	1418
δ scenario					
Mean deciview change	−12 dv	−12 dv	−9 dv	−9 dv	
Mean	$66	$91±$25	$84±$34	$57±$14	
Median	$23	$50	$11	$5	
Mean (5% trim)		$62±$18	$40±$15	$28±$6	
Respondents (*n*)	223	223	222	764	
Mean of predicted values		$72±$28	$55±$37	$12±$11	
Pooled respondents (*n*)		1215	1215	1215	

Note: Chestnut and Rowe deciview changes are based on the actual visibility in the Chestnut and Rowe photographs and, therefore, differ from the visibility changes reported by Chestnut and Rowe. Values from the non-Virginia sample of Chestnut and Rowe have been inflated to 2003 dollars using a factor of 1.55. Average deciview changes for 2003 survey are computed from the full distribution. Mean values from the 2003 survey show the 95% CI.

did so by envisaging a subject commodity larger than the one under investigation. That is, Chestnut and Rowe sought respondent estimates to help correct for any so-called "part/whole" bias. Chestnut and Rowe hypothesized that the respondent might be valuing general protection or improvement for national parks, rather than focusing on just the visibility issue, and asked questions intended to identify what fraction of the reported WTP was specifically for visibility. Despite concerns about how well respondents can be expected to provide meaningful quantitative answers to these debriefing questions, the authors closely followed them and Chestnut and Rowe's use of the responses to adjust the raw WTP estimates. The mean of Chestnut and Rowe's adjustments through this procedure was to 62% of the raw WTP estimates. In the survey of this study, the comparable statistic ranges from 78% to 85% for Versions 1 through 4, and is 61% for Version 5.

WTP Estimates from the New Survey

In this section, the authors present their empirical analysis of the WTP responses from the 2003 survey. They find evidence that: (1) WTP is highly skewed; (2) whereas systematic variation in WTP is difficult to explain using the demographics of respondents, WTP declines substantially and statistically significantly as respondents are reminded of budget constraints; (3) the public places a relatively high share of its value on reducing haze during the worst days relative to haze reduction during all summertime days; and (4) the public places value on avoiding a massive deterioration in visibility. All of these results contribute to the policy debate on regional haze.

Before turning to the specific results, it is useful to describe Table 5, because it summarizes all of the empirical results discussed in this section. Table 5 reports median WTP values, as well as untrimmed, trimmed, and predicted means for each version of the 2003 survey. It also reports comparable estimates from Chestnut and Rowe, along with the corrected visual range changes that accompany those WTP estimates. For policy purposes, it may be advisable to focus on the trimmed values to diminish the influence of outliers. For example, the National Oceanic and Atmospheric Administration Blue Ribbon Panel[7] made the case for such conservatism in this situation. Table 5 also reports the mean of predicted values using a standard multivariate regression technique that explicitly accounts for the censored nature of WTP.[16] In addition, to assess statistical significance, Table 5 also displays 95% confidence intervals.

We note that the results in Table 5 pass the standard internal and external scope tests, as discussed in Smith and Osborne,[14] which can be applied. For example, the results for the α and β scenarios for versions 3 and 4 are based on independent samples and are consistent with an external scope test, because the α WTP is less, on average, than the β WTP. The results for the β and γ scenarios for Version 5 are also based on independent samples and are consistent with an external scope for the same reason. Similarly, these results are consistent with an internal scope test, because the average β WTP is less than the γ WTP for all of the versions.

Distribution of Expressed WTP Values

Consistent with many CV studies, the authors find that the distribution of respondents' WTP is highly skewed. A large number of respondents expressed a very low WTP, whereas a small minority of respondents expressed relatively large values. This minority of large value responses considerably increases the mean WTP without affecting the median. As a result, the median and mean values of WTP shown in Table 5 differ substantially. These findings are consistent with Chestnut and Rowe. For example, for the β scenario of version 1, the untrimmed and trimmed mean WTP are \$77 and \$55, respectively, whereas the median is only \$35. For the γ scenario of version 1, these measures are \$118, \$70, and \$50, respectively. In all of the versions and scenarios, the authors find that the median WTP is never more than ~50% of the untrimmed mean and is typically only about a third of the untrimmed mean. The disparity between the mean and the median grows as respondents are reminded of budget constraints.

As a thought experiment to evaluate an implication of the skewed distribution of WTP, suppose that respondents were offered a public referendum on haze reduction in which the mean WTP was specified as the cost per person to reduce haze by amounts similar to those in the 2003 survey. For versions 3 and 4, which provide reminders of budget constraints, 70–80% of the respondents in this study would reject the mean value. Based on these responses, one would infer that the proportion of respondents who would reject that referendum is quite high.

Stronger Budget Reminders Reduce Average WTP

Multivariate regression analysis allows the authors to examine whether budget reminders reduce the average WTP by statistically significant amounts taking in account demographic attributes of respondents. However, WTP is censored at zero, and the statistical specification should explicitly reflect that feature. Following other WTP research, the authors use a Tobit specification to evaluate the effects of budget reminders.[17] For these analyses, the authors pooled the data across versions within a given scenario and included so-called "protest" bids. They used indicator (or dummy) variables for the version of questionnaire and included in these regression models measures on respondents' income, age, education, location, breathing problems, and prior visits to national parks. Providing the detailed regression results would substantially lengthen this paper; they, together with the programming code that generates them, are available from the authors on request. The signs of the coefficients are generally consistent with theory. For example, income has a positive effect on WTP. Whereas the coefficients are jointly significant, the authors find that the predictive power of the respondents' measures is low. As such, examining the simple trimmed means may be equally informative.

In Table 5, the authors report the average of the predicted WTP values from the Tobit specifications, with negative predicted values set to zero but included in the calculation of the average. To assess statistical significance, the authors use standard bootstrapping techniques with 1000 replications to calculate 95% CIs.[18] They find that the use of budget reminders reduces the average WTP by statistically significant amounts. For example, for the β scenario, the average WTP falls from $54 for version 2 to $33 for versions 3 and 4 to $9 for version 5. Based on the 95% CIs, the decline from $54 to $9 is statistically significant. For the γ scenario, the average WTP falls from $84 for version 2 to $34 for versions 3 and 4 to $14 for version 5. Again, based on the 95% CIs, the decline from $84 to $14 is statistically significant.

High-Valued Reductions in Haze

A limitation of the paper by Chestnut and Rowe was its inability to present visually those improvements that would have very little impact on average summertime conditions but would affect a limited number of days on which haze was particularly acute. Recognizing this limitation, the 2003 survey explicitly incorporates the α scenario. As described earlier, the α scenario essentially takes 9 of the 11 worst summer days and makes them moderately better, as contrasted with the other 2003 survey scenarios, which represented proportional shifts during all summertime days. Specifically, the α scenario amounts to a 1-dv improvement on average, compared with a 5-dv improvement for the β scenario and a 10-dv improvement for the γ scenario.

Despite the modest average improvement of the α scenario, the respondents in this study valued it at a relatively high proportion of the value they placed on the β scenario. For example, the untrimmed mean for the α scenario is $42 compared with $59 for the β scenario, a proportion of 71%. For the trimmed mean, these numbers are $21, $33, and 64%, respectively. For the average of the predicted values from the multivariate regression, these numbers are $16, $33, and 49%, respectively. When compared against the γ scenario, which represents an average change that is one order of magnitude larger than the α scenario, these proportions remain large. For example, the untrimmed mean for the α scenario is $42 compared with $76 for the γ scenario, a proportion of 55%. For the trimmed mean, these numbers are $21, $39, and 54%, respectively. For the average of the predicted values from the multivariate regression, these numbers are $16, $34, and 47%, respectively.

These findings are consistent with the hypothesis that the public places a relatively high value on improving visibility on the worst days, with relatively less value being placed on improving moderately clear days. They also reinforce the importance of using frequency distributions to describe visibility changes as they permit researchers to convey more specific information on the nature of the change being valued. Using the Chestnut and Rowe methodology to depict changes on average, a difference of 1 dv would hardly be perceptible and would presumably be valued very close to zero.

Value of Avoiding Massive Deterioration

The δ scenario in the 2003 survey represents a massive deterioration in visibility during summertime days. As such, it has inherently less relevance in policy-making than the scenarios that depict visibility improvement. The findings in this study on the valuation of avoiding deterioration are broadly consistent with those of Chestnut and Rowe. The authors find that, in all of the versions, the value placed on avoiding a massive deterioration is typically greater than or equal to the value placed on the α and β scenarios. For example, in version 1, the average of the predicted values from the multivariate regression is $72 for the δ scenario and $62 for the β scenario. This relationship does not always hold, however. The authors also find that the value placed on avoiding a massive deterioration is less than the value placed on the γ scenario.

ACKNOWLEDGMENTS

The authors wish to acknowledge Electrical Power Research Institute for the funding to support this research. The authors thank three anonymous reviewers for their useful comments and Mireille Bishay for her valuable research assistance.

REFERENCES

1. Mitchell, Robert Cameron; Carson, Richard T. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Resources for the Future: Washington, DC, 1989.
2. *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*; Bateman, I.J.; Willis, K.G., Eds.; Oxford University Press: New York, NY, 1999.

3. Chestnut, L.G.; Rowe, R.D. Preservation Values for Visibility Protection at the National Parks; Draft Final Report under U.S. EPA Cooperative Agreement No. CR813686; Prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Economic Analysis Branch: Research Triangle, NC, 1990.
4. Regulatory Impact Analyses for the Particulate Matter and Ozone National Ambient Air Quality Standards and Proposed Regional Haze Rule; Innovative Strategies and Economics Group, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency: Research Triangle Park, NC, 1997.
5. Regulatory Impact Analyses for the Final Regional Haze Rule; Innovative Strategies and Economics Group, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency: Research Triangle Park, NC, 1999.
6. The Benefits and Costs of the Clean Air Act 1990 to 2000*;* EPA 410-R-99–001; Prepared by U.S. Environmental Protection Agency for U.S. Congress, Office of Air and Radiation, Office of Policy: Washington, DC., 1999.
7. Kemp, M.A.; Maxwell, C. Exploring a Budget Context for Contingent Valuation Estimates. In *Contingent Valuation: A Critical Assessment*; Hausman, J., Ed. North-Holland Publishing Company: New York, 1993. pp. 217–269.
8. Report of the NOAA Panel on Contingent Valuation. *Fed. Regist.* **1993**, *58* (10), 4601–4614.
9. IMPROVE Home Page, available at http://vista.cira.colostate.edu/improve/ (accessed October 3, 2005).
10. Molenar, J. Warnings on the Use of Transmissometer Data to Analyze Long-Term Visibility Trends; available at http://vista.cira.colostate.edu/improve/Publications/GrayLit/TransDataUseWarning/TransDataUseWarning.htm (accessed October 3, 2005).
11. Reconstructing Light Extinction from Aerosol Measurements; available at http://vista.cira.colostate.edu/improve/Tools/ReconBext/reconBext.htm (accessed October 3, 2005).
12. IMPROVE Aerosol Data; available at http://vista.cira.colostate.edu/improve/Data/IMPROVE/AsciiData.aspx (accessed October 3, 2005).
13. Prepared by John Molenar of Air Resource Specialists: Fort Collins, CO.
14. Smith, K.V.; Osborne, L. Do Contingent Valuation Estimates Pass a "Scope" Test? A Meta-Analysis; *J. Env. Econ. Manage.* **1996**, *31,* 287-301.
15. Chestnut, L.G.; Dennis, R. Economic Benefits of Improvements in Visibility: Acid Rain Provisions of the 1990 Clean Air Act Amendments; *J. Air & Waste Manage. Assoc.* **1997**, *47,* 395-402.
16. Greene, W.H. *Econometric Analysis,* 4th ed.; Prentice-Hall: Upper Saddle River, NJ, 1999.
17. Halstead, J.M.; Lindsay, B.E.; Brown, C.M. Use of the Tobit Model in Contingent Valuation: Experimental Evidence from the Pemigewasset Wilderness Area; *J. Env. Manage.* **1991**, *33,* 79-89.
18. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Application*; Cambridge University Press: Cambridge, UK, 1997.

About the Authors

Anne E. Smith is an economist and a vice president; Michael A. Kemp is a full member of the market research society, the director of survey research, and a vice president; and Timothy H. Savage is an economist and a principal with CRA International. Catherine L. Taylor is now an associate at the Brattle Group. Address correspondence to: Anne Smith, CRA International, 1204 F St. NW, Ste 700, Washington, DC 20004; phone: +1-202-662-3872; fax: +1-202-662-3910; e-mail: asmith@crai.com.