Comments on the First External Review Draft of EPA's "Integrated Science Assessment for Particulate Matter"

Anne E. Smith, Ph. D.
CRA International
1201 F Street NW, Suite 700
Washington, DC 20004

Prepared at the request of the Utility Air Regulatory Group
March 13, 2009

In December 2008, the U.S. Environmental Protection Agency (EPA) released a first external review draft document titled "Integrated Science Assessment for Particulate Matter" (hereafter, the "Draft ISA"). Following are my comments on the Draft ISA's Section 9.3.5, concerning welfare effects from reduced visual air quality (VAQ).

1. Summary of Main Points

1. New research by CRA summarized in my comments indicates that VAQ preference studies are not a reliable method for eliciting opinions about a cutoff for acceptable VAQ, and therefore should not be used in setting any visibility standard. The level of the cutoff or "standard" that they indicate appears to be strongly determined by the range of VAQ used in the survey. The Draft ISA should address the new CRA results in its review of the VAQ preference studies.

2. The Draft ISA appropriately identifies separation of health concerns from aesthetic concerns as a primary concern with VAQ valuation and preference studies. New evidence from CRA research reinforces this concern, showing that many of the respondents to both types of surveys intertwine health risk concerns with aesthetic concerns, even when instructed to avoid doing so. The Draft ISA should incorporate the additional new evidence described herein.

3. The Draft ISA identifies only one new visibility valuation study, a property value study by Beron *et al*. (2001), and is appropriately cautious in noting concerns that its value estimates might be contaminated by intertwined health risk values. I elaborate on the reasons for those concerns, and why Beron *et al.* (2001) should not be used in setting a visibility standard.

4. The Draft ISA fails to identify a new visibility valuation study published since the last PM Criteria Document. Smith *et al.* (2005, 2006) report on a major new contingent valuation study of visibility value that includes residential use value as well as recreational and preservation value. The Draft ISA Section 9.3.5 should review this study because it provides important new insights about the interpretation and sensitivity of all visibility contingent valuation studies.

A number of corrections also are needed to several statements made in Section 9.3.5 of the Draft ISA. These are identified in the following comments as well.

2. Comments on VAQ Preference Studies

"Urban visibility preference studies examine individuals' preference [*sic*] by investigating the basic question 'what level of visibility degradation is unacceptable'."[1] These are not economic studies in that they purposely avoid considering what people might be willing to pay in order to attain the VAQ that they say they prefer. Ely *et al*. (1991) is the original VAQ preference study, and all the others identified in the Draft ISA follow the format of this paper. Ely *et al*. (1991) start with the following view on how to set a visibility standard:

> "[A]n air quality standard must formally state, without regard to possible cost of attainment, how much of a pollutant or what effects of pollutant(s) are unacceptable. A visibility standard should reflect *enduring judgments* regarding the appearance of the air rather than how much visibility improvement can be afforded at a particular point in time. The level chosen needs to be credible, and consistent with what the public's own eyes tell them is the cutoff between acceptable and unacceptable VAQ."[2]

Consistent with this view, the VAQ preference studies are surveys that show a sample of local residents multiple photographs of different levels of visibility in their city, and ask them which of the photographs represent "unacceptable" conditions. They are told to think only of the visibility itself when answering this question. The studies interpret the visual range below which more than 50% of the sample deem VAQ to be "unacceptable" as the visual range at which the visibility standard should be set for that city.

In addition to the Ely *et al*. (1991) VAQ preference study, which was for Denver, the Draft ISA identifies two other full VAQ preference studies and one smaller-sample "pilot" study that repeated the general approach of Ely *et al*. in different locations. The full studies are for Phoenix (BBC Research & Consulting, 2003) and for two suburban locations in British Columbia (Pryor, 1996). The pilot study was for Washington D.C. (Abt Associates, 2001).

In regard to these four VAQ preference studies, the Draft ISA concludes: "One notable finding is the general degree of consistency in the median preferences for an acceptable level of visibility degradation," which it reports to be in a range is 19 to 25 deciviews (dv), or 30 to 55 km.[3] This range is not correct. The correct range is 19 to 29 dv, or 21 to

[1] Draft ISA, p. 9-4.
[2] Ely *et al*. (1991), p.3, emphasis added.
[3] Draft ISA, p. 9-4.

55 km.[4] The Draft ISA should be revised to report the correct range. Further, this range hardly reflects a "general consistency." The upper end of the range has 185% more light extinction (b_{ext}) than the lower end of the range. The range of $PM_{2.5}$ concentrations that would be associated with this range of potential VAQ standards could be even larger. For example, a standard of 19 dv could imply a $PM_{2.5}$ limit as low as about 7.5 $\mu g/m^3$ in some areas, while a standard of 29 dv could imply a $PM_{2.5}$ limit as high as 46 $\mu g/m^3$ in some areas.[5] This hardly appears to be "general consistency" in median preference levels for VAQ. The Draft ISA should delete the characterization of "general consistency," state the range correctly, and provide context for how large this range is from the perspective of stringency of a $PM_{2.5}$ standard set at its upper or lower bound.

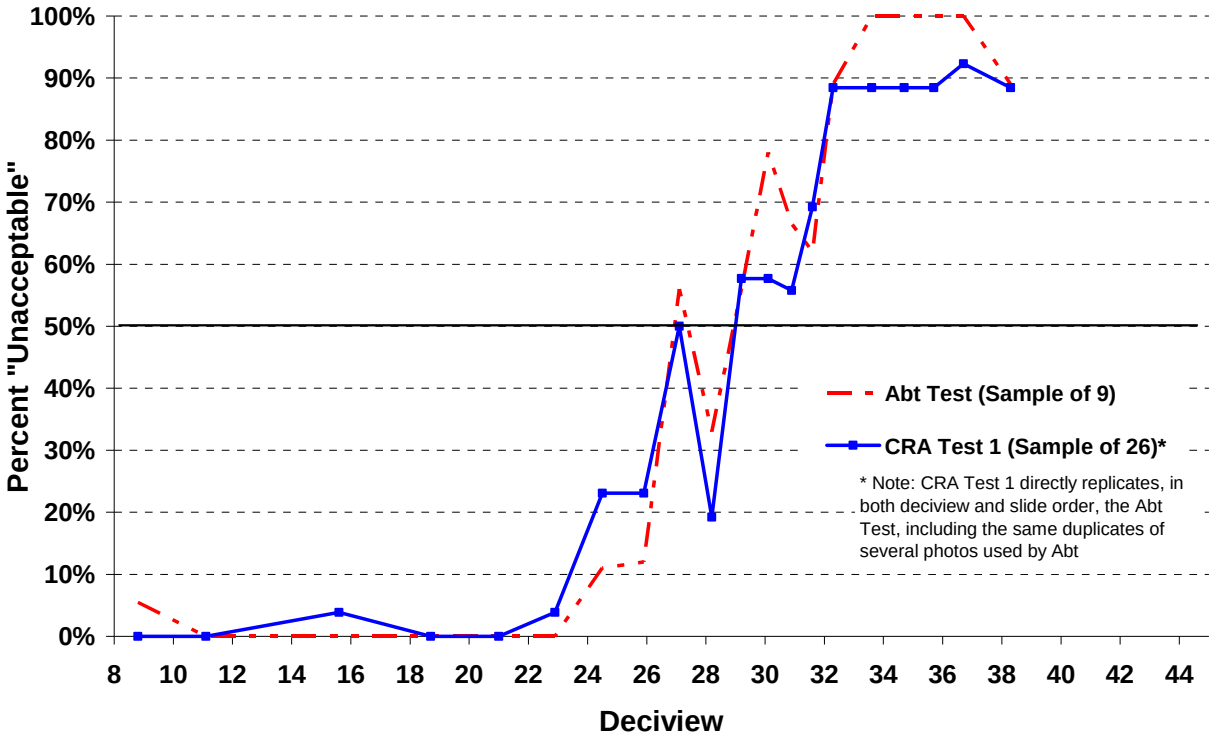
Consistency does exist, however, in the fact that almost all respondents consider at least one or more of the photographs to have "unacceptable" VAQ. Could it be possible that people are likely to identify at least one VAQ level as unacceptable when they are asked to decide if each of a set of images with varying VAQ levels is acceptable or unacceptable? To test this question, CRA conducted a quick study with a sample of 60 people, and found that the range of VAQ shown in a survey readily alters what the survey will "find" as the cutoff point of acceptability. Thus, the cutoff for acceptability is neither an "enduring judgment," nor credible, as Ely *et al*. called for. Rather, the cutoff depends on the range of VAQ that the survey designers decide to use. This finding clearly undermines the credibility of VAQ preference studies categorically, but it also poses a serious question of whether any absolute notion of "unacceptable VAQ" even exists. Certainly, these four VAQ preference studies do not offer a credible or enduring indication of what a visibility "standard" ought to be and should not be used for that purpose.

[4] EPA appears to have mis-read the results of Abt Associates (2001) for Washington DC. That survey finds that photographs varying from 27 to 31 dv were deemed "unacceptable", as can be seen in Exhibit 8 on p. 16 of Abt Associates (2001). In that Exhibit, the 50% level is achieved for photographs between 20 and 30 $\mu g/m^3$, which (per Exhibit 2, p. 5) is 27.1 to 30.9 dv. The point where the percent "unacceptable" remains above 45% for all worse VAQ photographs is 29.2 dv. This interpretation of Exhibit 8 is corroborated by the report's text, which states "The slides with concentration levels of 25 $\mu g/m^3$ and 30 $\mu g/m^3$ are considered the 'threshold' slides. Unacceptable/acceptable rates were divided for these slides, and a few participants noted in their response booklet that the visibility was 'so-so,' 'border line' and a 'middle choice'." (Abt Associates, 2001, p. 18). The 25 $\mu g/m^3$and 30 $\mu g/m^3$slides were 29.2 dv and 30.9 dv, respectively. This error in how the Draft ISA describes the cutoff in Abt Associates (2001) appears in Table 9-2 (p. 9-77) and on p. 9-78 (line 5), as well as on p. 9-4; it should be corrected in all three locations.

[5] Using the traditional IMPROVE algorithm for estimating light extinction on p. 9-15 of the Draft ISA, 19 dv can occur with a PM mix of only 5, 0.5, and 2 $\mu g/m^3$ for sulfate, nitrate, and organic carbon, respectively, if the relative humidity is 75% and the organics mostly hygroscopic. 29 dv can occur for a PM mix of 3, 0.2, 8, 10, and 25 $\mu g/m^3$ for sulfate, nitrate, organic carbon, elemental carbon, and fine soil, respectively, if the relative humidity less than 40% and the organics not hygroscopic. (Even when restricting measurements of visibility to just the hours of noon to 4 p.m., average relative humidity stays consistently below 40% in some parts of the country, yet can frequently rise as high as 75% in other parts of the country.)

CRA's experiment was to recreate the surveys done previously, using the Abt Associates (2001) pilot as the model. A sample of 26 individuals were shown photographs of a vista of Washington D.C. (also used by Abt Associates), with the exact same VAQs, ordering and duplicates that Abt Associates used.[6] This was intended to establish a baseline, as well as to show that our quick survey could reproduce the results found by Abt Associates. Figure 1 compares the results from this CRA sample ("Test 1") with those reported in Abt Associates. It is clear that differences of venue, sample selection, and the presentation of the photographs do not undermine our ability to use our own survey to test the robustness of results to survey design. In fact, even the irregularities in the monotonicity of the ratings found by Abt Associates are replicated in the CRA results.[7]

Figure 1. Comparison of Abt Associates (2001) and CRA Results on Percent "Unacceptable" by Deciview (Abt results from Abt's Exhibit 11, where scores for slides that were rated twice were aggregated to a single score for that VAQ level)

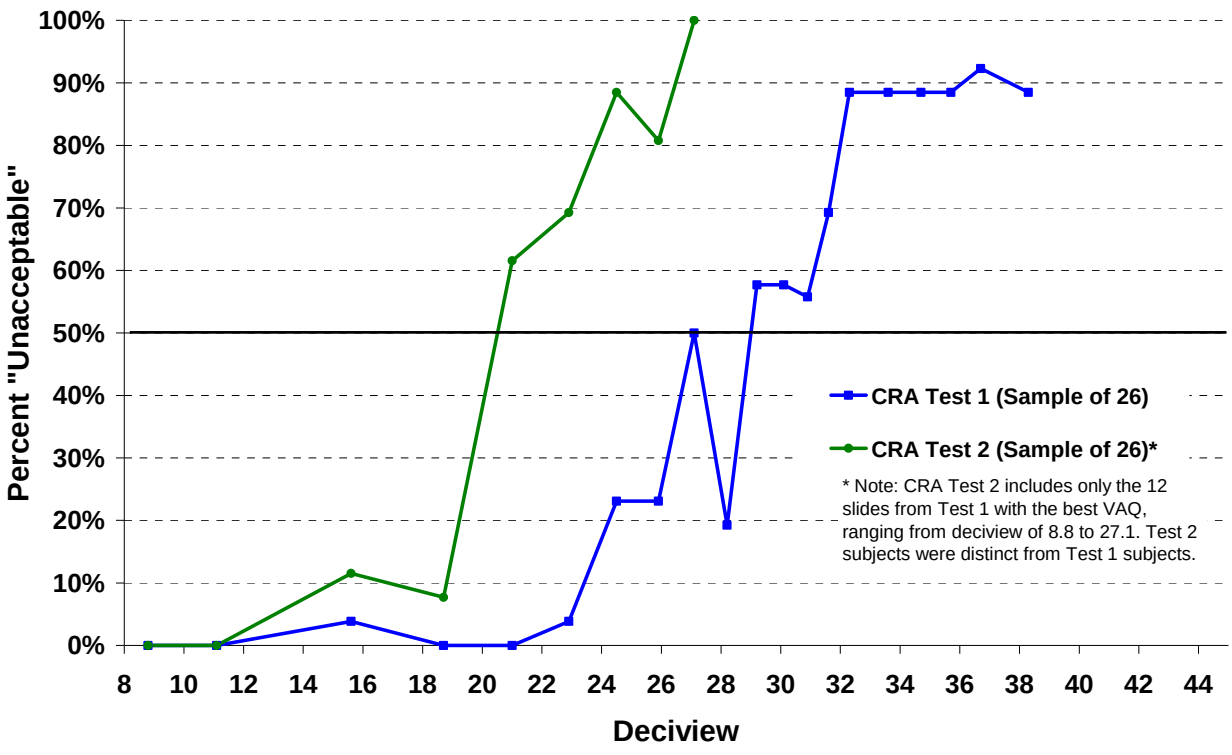


[6] These were documented in Exhibit 2, on p. 5 and in Appendix A on p. 50, respectively, of Abt Associates (2001).

[7] For example, both samples produced a pronounced dip in the ratings at 28.2 dv. We hypothesize this dip can be explained by the contrast between that slide and the one that immediately preceded it in both the surveys.

CRA also subjected a different set of 26 individuals to the same exercise, except that the worst VAQ shown to this sample group was only 27.1 dv, which was the lowest VAQ that had been deemed unacceptable by 50% of the Abt sample ("Test 2").[8] If the original survey were able to indicate an absolute point at which the median of the population considers VAQ to be unacceptable, then the group taking Test 2 should not have assigned more than about 50% unacceptability to even the worst of the VAQ conditions that they were shown. However, the Test 2 group *also determined that there was a cutoff point well within the range of VAQs that they were shown* – and the dv cutoff from Test 2 was dramatically lower than the dv cutoff from Test 1. The results for the two samples are presented in Figure 2. (The line for Test 1 in Figure 2 is identical to the line shown in Figure 1 for the CRA replication of the Abt survey.) Figure 2 provides strong evidence respondents have a natural inclination to find at least some photographs "unacceptable" when asked to rate a set of them as either acceptable or unacceptable.

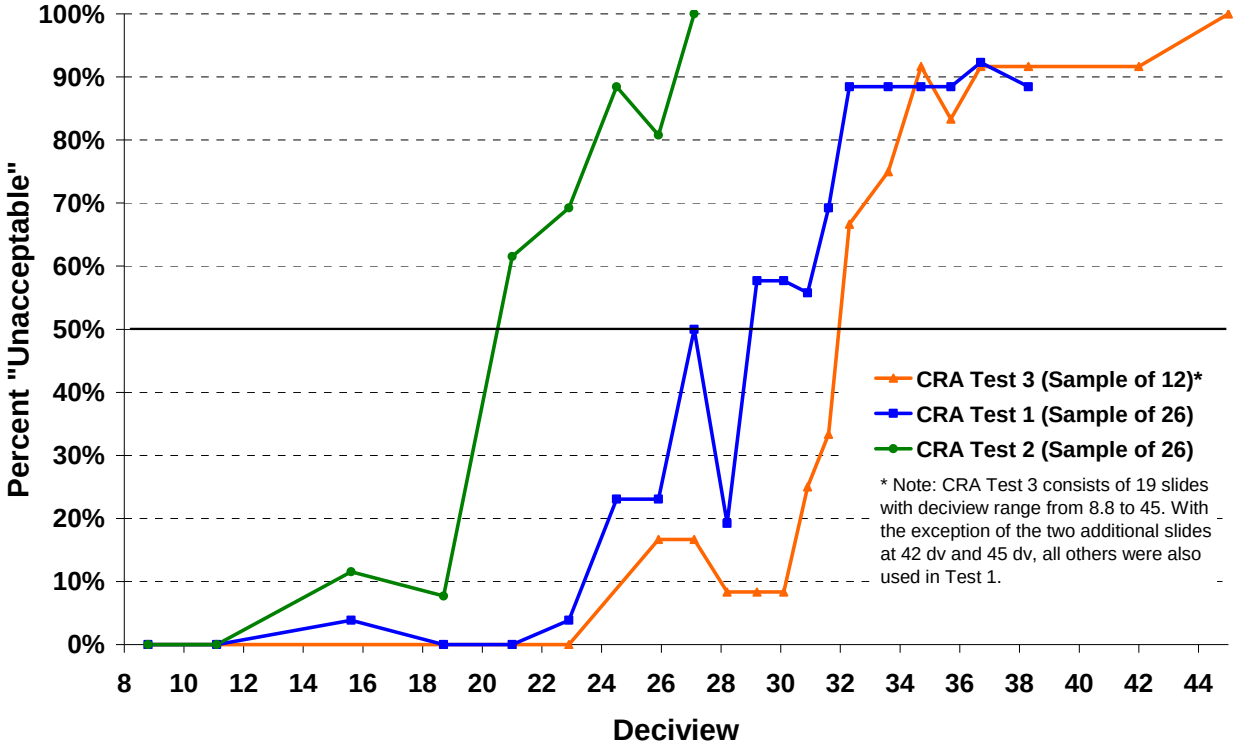
Figure 2. CRA Results on Percent "Unacceptable" by Deciview for "Test 2" (Showing Only Photographs Up to 27.1 dv) Compared to Results for "Test 1" (Showing Photographs up to 38.3 dv, as in Original Abt Survey)



[8] Individuals were randomly-assigned to test versions.

If the median cutoff VAQ is dependent on the range of VAQs viewed by a sample of people, this effect should also appear in the upward (less stringent cutoff) direction if a wider range of VAQ is shown. To check, CRA created a "Test 3" that included two photographs with higher deciviews (*i.e.*, worse visibility) than the highest in the Abt survey, and included these in the survey given to a third group of 12 individuals.[9] Figure 3 adds the results for "Test 3," which also are consistent with the hypothesis that the acceptability cutoff in VAQ preference surveys is contingent on the range of visibility levels included in the survey.

Figure 3. Results for "Test 3" (Showing Photographs up to 45 dv) Compared to Tests Showing Less Range in Potential VAQ



CRA's study provides strong evidence that VAQ preference studies are unreliable as a basis for setting a standard for visibility, as no absolute concept of "unacceptable VAQ" appears to exist. The CRA study shows that previous VAQ preference surveys cited in the Draft ISA do not offer a scientifically credible method for identifying a visibility standard – not even on a city by city basis.[10]

[9] Compared to a maximum of 38.3 dv in the Abt survey, photographs in CRA's Test 3 also included photographs of 42 dv and 45 dv.

[10] In fact, it also suggests that the can be no such thing as a uniform national standard for visibility. If the photographs for a preference survey in each location of the country were to be carefully tailored to show VAQ ranges currently experienced in each area, the "standard" would end up varying by location.

Estimates of a VAQ acceptability cutoff are probably also dependent on the particular base photograph selected for use in the survey. Only one of the four preference studies offers any possibility of observing such an effect. Pryor (1996) asked his respondents to identify the acceptability of VAQ for two different British Columbia vistas from within the Lower Fraser Valley. Both included mountain ranges in the far-field, but one had diverse urban structures in the mid-field and the other had a sparser, more uniform set of buildings in the mid-field. Each respondent was asked to rate the acceptability of VAQ for both vistas. The cutoff identified for the VAQ in this same general area was 18.6 dv for one vista and 23 dv for the other. Clearly the choice of base photograph can also strongly affect what these VAQ preference studies say a visibility standard should be, if they were to be used to set such a standard.

The evidence from CRA's experiment with the Washington DC vista combined with this sensitivity to the choice of vista found by Pryor (1996) indicate that VAQ preference studies do not offer a valid basis for identifying a visibility standard – either nationally or locally. They do not measure any "enduring judgment," but a judgment that varies with the way alternative VAQ conditions are presented to the respondent. The degree to which that judgment varies with survey design is too wide to offer any information on how to set a VAQ standard. The main conclusion one can draw from these preference studies is that people are able to discern varying degrees of VAQ, and that they do prefer clearer air to less clear air. These facts, however, have already been established in earlier research.

There is another problem with the VAQ preference surveys that is a serious concern for the visibility valuation studies as well. A substantial number of respondents incorporate into their responses about willingness to pay for aesthetics their personal beliefs that poor visibility due to pollution also implies health risks from breathing the pollution. The Draft ISA highlights this problem as its key concern in using any of these studies:

> "The degree to which previous human preference and valuation studies for VAQ have adequately made this distinction and separation [between the aesthetic and wellbeing components associated with the visibility condition] is an important issue in applying results from available studies in a Secondary NAAQS"[11]

Some might defend the preference studies because they instructed respondents not to consider health impacts when answering the survey questions about acceptability. However, when follow-up questions have been asked, many respondents have said they made their judgments by considering how unhealthful the air must be if the visibility is poor. Among the VAQ preference studies, only the Abt survey and CRA's recent replication of the Abt survey asked any follow-up questions about what respondents had taken into consideration when thinking about "acceptability" of the VAQ in the

[11] Draft ISA, p. 9-75.

photographs. In both studies, some of the respondents mentioned health risks.[12] For example:

> "We also asked participants if they were thinking about the negative health effects that might be associated with the higher level of visibility impairment in some slides. One participant is employed at an environmental organization, and found it difficult not to associate the visibility conditions with adverse environmental and public health effects. Other participants discussed that it was difficult because they know that air pollution and negative health effects 'go together.' For example, one participant discussed that the level of visibility impairment shown in the slide made her think of the negative health effects associated with a day similar to the one shown. She discussed that on a day similar to the one shown, her sinuses would bother her, and she would feel that the air quality for the day was bad for her health. Another participant discussed that he works outdoors and that while viewing the slides he was thinking about how long his crew could work in the visibility conditions shown."[13]

Follow-up questions in the Abt Associates (2001) survey also found that respondents may intertwine feelings about weather conditions that they associate with the poorer visibility conditions (*e.g.*, high humidity), even though they were also told that there was no difference in the weather from photograph to photograph.[14]

The Draft ISA provides a misleading synopsis of the nature of health confounding in the Abt Associates study in the following excerpt:

> "1. Participants had been asked how they reacted to the initial direction to base their answers only on visibility, but health was never explicitly mentioned by the focus group moderator. Participants strongly agreed with the decision to not mention that health effects are associated with visibility impairment. They understood the directions as meaning they

[12] The Draft ISA in Table 9-2 (p. 9-77) says that the Abt survey never mentioned health during the survey. This is not correct. The script of that survey contains the following statement just before asking respondents to rate the photographs: "You may also know that the same air pollutants that form haze and reduce visibility are of concern because they are harmful to people and the environment. For this discussion, we want you to focus only on visibility: how well you can see. Our goal is to understand your opinions regarding visibility itself. It is important that you focus on how things look to you in the slides you are going to see. When you give us your opinion, consider only the visibility. Just focus on how things look: how far you can see; the clarity; the crispness; the colors." (Abt Associates, 2001, p. 59). Abt Associates says that the "script does not discuss the health risks" (p. 27), but the script itself shows that health effect were indeed mentioned – albeit not "discussed." Table 9-2 of the Draft ISA should be corrected to note that health effects were mentioned with the words "harmful to people" when telling respondents to consider only visibility in their responses.

[13] Abt Associates (2001), p. 18.

[14] Abt Associates (2001), p. 18.

> should ignore health issues, and said their answers would have been different if they included health as well as visibility in their judgments."[15]

The first sentence is not really consistent with the script itself (see footnote 12), in which "health" is not mentioned, but "harmful to people" *is* mentioned. Despite some statements by the Abt authors themselves that health was not mentioned, Abt's official script can be understood otherwise.[16] Second, while one participant did say his answer would have been different if he had "known that the level of visibility impairment would affect his health,"[17] this is not the main point in what Abt reports about health effects values. The main point is that people *did* intertwine health effects concerns with aesthetic concerns, even though they were told to think only about the visibility, and not about how the air might be "harmful to people and the environment."[18] The following statement in the Abt report is the more important conclusion that the Draft ISA should highlight instead:

> "During the evaluation of our pilot focus group sessions, we asked participants if they thought they were successful in separating the two issues [*i.e.*, health concerns and visibility conditions]. Responses were split, with some participants saying they could make the distinction, while others replied that health concerns were factored into their decision making."[19]

The quote from p. 9-81 (lines 14-19) of the Draft ISA therefore should be replaced with the finding that the Abt study's follow-up questions did identify problems on the part of at least a portion of the respondents in separating health from aesthetics in their judgments. CRA found the same phenomenon in its replication of the Abt survey. People understood they were supposed to consider only aesthetics, but a portion of the respondents considered it anyway. Given that none of the three other preference studies asked such a follow-up question, there is no way infer that these studies' results are not similarly confounded by health values. However, it would be inappropriate to suggest that they are less likely to be affected by this problem because their scripts mentioned health and the Abt script did not. All the scripts did mention health, albeit not always using that exact word, and all scripts mentioned it only in order to tell respondents not to consider health in their responses. If anything, the fact that one respondent in the Abt survey said he would have changed his answer if he "knew" the visibility was also harmful to health suggests that the more direct mention of health effects in the other surveys might have *exacerbated* the degree to which health concerns may confound those studies' results.

[15] Draft ISA, p. 9-81, lines 14-19.
[16] Nothing in the Abt report suggested that the script was not followed exactly as documented.
[17] Abt Associates (2001), p. 27.
[18] Abt Associates (2001), p. 59.
[19] Abt Associates (2001), p. 27.

Thus, the Draft ISA is correct to have concerns about the entanglement of health values with aesthetic values in the preference studies as well as in the valuation studies. However, the inability of the VAQ preference study approach to identify any absolute concept of "unacceptable" is a far greater concern, because it means that this technique simply should not be used in setting standards for visibility.

3. Comments on Visibility Valuation Studies

3A. Critique of Beron *et al.* (2001)

The Draft ISA also discusses urban visibility valuation studies (or "economic studies"), which it describes as studies that "examine preference by investigating 'how much would you be willing to pay to improve visibility'."[20] Six of the seven urban visibility valuation studies cited in the Draft ISA are 18 to 30 years old, and were accounted for in the 2004-2005 PM NAAQS review.[21] Beron *et al.* (2001) is the one "new" study that the Draft ISA identifies, although it also pre-dates the 2004-2005 review. (EPA failed to identify it during the previous review cycle.) For reasons explained below, the Draft ISA is correct to treat this one new study with great caution.

All but one of the six other urban visibility valuation papers cited in the last PM NAAQS review are based on studies that directly asked individuals about their willingness to pay for altered VAQ conditions, represented with photographs. (In economists' jargon, this is the "contingent valuation" method, also called the "stated preference" approach.) The sixth of the older studies employed a valuation technique that statistically estimates the determinants of residential property values, including a variable for the visual range at each home sold during a study period. (In economists' jargon, this is the "hedonic" method, also called the "revealed preference" approach.) The hedonic method has long been deemed unreliable for valuing the aesthetic aspects of visibility changes because one cannot "instruct" homebuyers to only consider the aesthetic aspects of the visual range that they see, while one can at least try to do that in a contingent valuation survey.[22] For this reason, Chestnut and Dennis (1997) concluded that the implied value of visibility in the single hedonic study that they reviewed was much higher than those obtained in the contingent valuation studies because "this is consistent with the expectation that the property value results will reflect values for all aspects of air quality, including concerns about health as well as visibility."[23]

[20] Draft ISA, p. 9-4.
[21] The ISA does not even cite these six studies directly, but only cites them by reference to a paper (Chestnut and Dennis, 1997) that reviews them. The six studies referenced in Chestnut and Dennis (1997) date from 1979 through 1991. None were published in a peer-reviewed journal, and one remains in draft form to this day.
[22] As will be discussed later, contingent valuation studies also appear to largely fail at eliminating health concerns from the values that they elicit from people, even though they at least have an opportunity to try.
[23] Chestnut and Dennis, (1997), p. 398.

The "new" study, Beron *et al.* (2001), uses the hedonic method that is widely recognized as problematic for valuing visibility, and likely to overstate the *aesthetic* values associated with visibility. Thus, it should be unsurprising that Beron *et al.* reports a very much larger value for visibility. The Draft ISA does not report, however, that Beron *et al.* is actually a direct descendent of the single hedonic valuation study reviewed by Chestnut and Dennis (1997), as described in the preceding paragraph (*i.e.*, Trijonis *et al.*, 1985). Two of the four authors are the same in both studies, the statistical formulas used in the 2001 paper are very similar to those in the1984 report, and both are for data from the Los Angeles area. Thus, it is useful to provide more information about what Chestnut and Dennis (1997) concluded regarding Trijonis *et al.* (1985).

Chestnut and Dennis (1997) pointed out that the most technically advanced of the contingent valuation studies (*i.e.*, McClelland *et al.*, 1993) explicitly explored how much of an initial stated value for visibility improvement might actually reflect concerns for health. They found that the value due only to visibility aesthetics was *less than one-fifth* the stated value for the changes in visibility shown in the survey. Respondents reported that the majority of the value was due to their concerns with health effects that they thought about when looking at photographs of poorer visibility conditions.[24] Taking into account information such as this, and their judgments on the quality of each study's design, Chestnut and Dennis provided their professional judgment of the best estimate for the value of a 20% change in visual range based all six of the studies that they were reviewing. The values in Trijonis *et al.* were higher 8 to 23 times higher than their opinion on a "best estimate." Chestnut and Dennis also suggested a range of uncertainty on the value for urban air visual aesthetics, based on all six of the studies they reviewed.[25] The range of values based on the hedonic property value method by Trijonis *et al.* was 6 to 16 times higher than the *upper* bound of Chestnut and Dennis's range of uncertainty.

Beron *et al*. (2001) has estimated values per increment of improved visual range that are up to five times higher than even the upper bound of the earlier Trijonis *et al.* study.[26] Beron *et al*., noting this large increase relative to the earlier Los Angeles hedonic study, conclude that the Trijonis *et al.* paper "understated the value of visibility

[24] Specifically, on p. 398-399, Chestnut and Dennis (1997) report that the fully-adjusted annual household WTP for a 20% change in visual range in McClelland *et al.* was $160 when before they controlled for concerns with health effects implicitly assumed to be associated with poorer visibility, and $29 after attempting to remove the portion of WTP attributable to health effects concerns. This is a factor of 5.5 difference.

[25] Stated in terms of WTP for a 20% change in visual range, the range of uncertainty they select, based on their professional judgment, is $22 to $41 (1994$), even though the range of estimates among all the contingent valuation studies in their review is $22 to $242, and the estimates from the hedonic study in their review range from $245 and $657.

[26] The Draft ISA reports the values found in the six earlier valuation studies and the Beron *et al.* study on a per Deciview basis, rather than per 20% visual range increment. To reconcile the values reported in the Draft ISA, p. 9-83, with those in the preceding footnote, which what one will find if actually reading the original Chestnut and Dennis paper, divide the values in the preceding footnote by 1.823, which is the number of deciviews of change per 20% change in visual range.

improvements."[27] However, they do not explain why one might not conclude instead that *their* paper is *overstating* the value of visibility. The rest of this section explains the technical basis for such an alternative conclusion.

Beron *et al*. employed a more sophisticated methodology with more extensive data than Trijonis *et al.*, and they included air pollutants (*i.e.*, ozone and TSP) in their regression along with visual range, which Trijonis *et al.* did not. They argue that this means that they have successfully controlled for the health-related concerns, and thus their estimate of the value associated with visual range now reflects aesthetic concerns only. But these facts alone do not enable them to conclude that they have produced better estimates of the value people attach to the aesthetic aspects of visual range changes. In order to have confidence that their new estimates reflect only aesthetic values, one must also offer a convincing argument that the following two conditions also hold:

1. That the purchasers of homes were not relying on their observations of visual range at the property to infer other value-enhancing attributes such as concerns about health effects from poor air quality.

2. That the homebuyers did know the relative ozone and TSP levels in each individual neighborhood.

There are strong reasons why the above two conditions do *not* hold for Beron *et al.*

Beron *et al*. believe that they have controlled for homebuyers' concerns about the healthfulness of the local air because when they include TSP and ozone in the regressions, those variables have an effect on home price that is independent of the effect of visual range, correctly signed, and statistically significant. From that fact, they conclude that the coefficient on visual range no longer includes any or much value associated with concerns about health effects that are associated with air pollutants. Beron *et al*. are failing to be introspective about the nature of the variables that they are using, and their relationship to what homebuyers can readily observe about a property. Assuming that homeowners do incorporate concerns about healthfulness of the air into their home purchasing decisions, they will rely on the information that is most readily observable to them about local air quality. Rightly or wrongly, what they can "see" at the property (*i.e.*, the visibility condition) is readily and directly observable. All of the visibility surveys that have included relevant follow-up questions have found that people automatically consider visibility to be a strong indicator of concerns about health effects.[28] In order for the hedonic method to be able to remove their values for healthful air from their values for visual range, a variable must be added that homebuyers also can observe and which indicates air pollution levels better than "what their own eyes tell them."

[27] Beron *et al*. (2001), pp. 334-5.
[28] See, for example, Chestnut and Dennis (1997), McClelland *et al*. (1993), and Smith *et al.* (2005, 2006).

The measures for TSP and ozone that Beron *et al*. employed do not have the observability necessary to compete with the directly observed visual range conditions. Although some measures of air quality are reported to the public, they are never offered with such locational specificity that a typical homebuyer would be able to assess whether a home in one neighborhood has generally higher or lower ozone concentrations than a home in another neighborhood. In fact, even the researchers had to construct that information with an advanced geostatistical technique known as kriging. The actual ozone and TSP data are only available for about 40 spots around the entire LA basin. The average homebuyer is very unlikely to have, or to seek out, information on the relative air pollution levels at these monitoring stations. They also have little incentive to do so because they tend to assume that they can assess the air quality from the visibility conditions that they do observe at every single individual location. They most certainly could not have performed the geospatial inferences about local air pollution levels from these 40 sets of monitoring data that Beron *et al.* rely on.

Thus, the act of including TSP and ozone in a hedonic study probably does nothing to account for buyers' valuation of healthfulness of air. The visibility coefficients (assuming the kriging provides a reasonable approximation of the actual local visibility that they observe) would still incorporate residents' health-related values, as well as their values for the aesthetic aspects of visibility. This confounding of the visibility estimate with health values will be the case *even if the buyers have a mistaken understanding of the actual air pollution levels*.[29]

The authors fail to think about this problem that their "explanatory" variables were not observable to the homebuyers. They appear to have confused availability of data to themselves (data that they constructed for themselves only) with knowledge of these data by buyers at the time of their purchase decisions. As a result, they come to an unjustifiable conclusion that the value attributable to ozone and TSP are for concerns with health, and that the value attributable to visual range therefore *solely* reflects aesthetic concerns. Their statistical techniques are highly sophisticated and appear to be thoughtfully conducted, but their interpretation of their results is highly flawed.

One might counter that a correctly-signed and statistically significant relationship between house price and ozone or TSP should not have been found if the buyers have no sound information on actual local air pollution levels. However, this could happen if some other undesirable attribute that is highly correlated with TSP or ozone might be observable in a neighborhood. In the case of ozone, one might consider the fact that high levels of ozone can damage materials. Conceivably, areas with higher average ozone levels may have experienced a more rapid degradation or fading of materials such as awnings, commercial banners, flags, or painted surfaces. High levels of ozone can also

[29] For example, the fact that ozone and TSP are not highly correlated with visual range in this study suggests that visual range observed at a location is not a very good proxy for actual health risk levels. Nevertheless, it is what the average buyer *believes* is the health risk and not what the health risk actually is that affects their willingness to pay for a home, which is all that a hedonic study estimates.

damage vegetation. Conceivably, the vegetation in a neighborhood that was experiencing higher average ozone could have been less aesthetic, possibly looking spotty or failing to thrive as lushly as in other neighborhoods. In short, higher ozone could possibly have been associated with a generally more dingy appearance in neighborhoods of similar affluence. If so, this could have created some distinction in home values that would be independently associated with kriged estimates of local average ozone levels. If this were occurring, however, the "value" attributed to the ozone coefficient in the hedonic equation would reflect values for *aesthetic* impacts of the ozone, rather than health-related concerns.

A similar possibility could apply to the measures of TSP. High levels of TSP also can have an aesthetic impact. Walls of buildings and concrete structures such as overpasses can become darkened if exposed chronically to high TSP levels. High localized TSP can also leave a dusting of grit and grime because much of it may settle locally. If TSP is very high, window sills, park benches and cars may therefore tend to be dirtier and windows to sparkle less, unless the neighbors and community wash these items more frequently than is the norm in the lower-TSP areas. These potential effects from higher TSP would be observable to homebuyers, and could affect neighborhood home sales prices relative to comparable homes in lower-TSP neighborhoods. If this were happening, the "value" attributed to the TSP coefficient in the hedonic equation would reflect aesthetic (i.e., "griminess") concerns.

Higher local TSP levels also could conceivably be an indicator for some other negative attributes of a home site that were not included in the Beron *et al.* regressions. TSP could be an indicator of proximity to a busy roadway, or proximity to industrial areas. Since the regression does not control for these either, the TSP-related value also could be reflecting willingness to pay to avoid living near undesirable activities or facilities near the home, or perhaps to avoid higher average noise levels. Under any of these possibilities, the TSP variable also would not necessarily have any relationship to health concerns associated with air pollution.

It would be useful to compare the values that Beron *et al.* find for TSP and ozone compare to those they find for visibility. If the TSP and ozone values reflect concerns with health, while the visibility value from the same regression reflects only aesthetics, one would expect substantially higher values to be estimated for TSP and ozone changes than for visibility changes. If not, that would lead to concerns that the study is not picking up the effects that the authors assert. In fact, if the TSP and ozone values are generally smaller than the visibility values, this would support the alternative and more likely interpretation of these values based on the fundamentals of what is observable (*i.e.*, that the hedonic value associated with visual range reflects buyers' willingness to pay for healthful air more than do the values associated with TSP and ozone).

The authors do not offer these highly-relevant comparisons of values, but I have estimated them from information in their paper.[30] My calculations indicate that the willingness to pay for a simultaneous 20% improvement in *both* ozone *and* TSP is only about 75% of the willingness to pay being attributed to a 20% improvement in visual range. This is consistent with the hypothesis that the ozone and TSP coefficients are reflecting non-health concerns while the visibility coefficient is capturing the bulk of health-related concerns as well as its own aesthetic concerns. It is *not* consistent with the authors' conclusions that the visibility coefficient reflects only aesthetic-related values while the pollutant variables have captured the health-related values.

The paper reveals another indication that the estimated association between visual range and home value may not reflect solely aesthetic attributes. The dummy variable for whether the property has a view ("VIEW") seems to have very little effect compared to the visual range. When dropped from the regression, the visual range coefficient only increases by 1-2 percent.[31] This raises the question of whether the visibility coefficient is truly reflective of an aesthetic value that can only be experienced if there is also a view. One explanation might be that the presence of good views around the neighborhood at large is more important to sales price than having a good view from the property itself. However, if the visibility values from this study are predominantly attributable to aesthetic values, then estimates of visibility values should at least be larger for homes that have a view of their own than for homes where the views are enjoyed only when traveling through the neighborhood. This could have been checked by incorporating an interaction term that would provide a separate value for visibility for homes with views and for homes without views. Given that the authors' primary interest in doing this study was to estimate visibility benefits, it is quite remarkable that they never considered this alternative specification, particularly when the VIEW variable on its own appeared to have no effect on the visibility value estimate. If an interaction term were not to be positive and statistically significant, it would be very difficult to assert that the visibility value being measured in this study is predominantly attributable to aesthetics and not to perceptions of health risk. Failure to check this obvious alternative specification is a significant omission in the study.

Thus, in direct opposition to the authors' conclusions, their estimates of how home values change with differing levels of visual range may still be reflecting peoples' concerns that poor VAQ means higher health risk. A portion of the VAQ-related value probably is attributable to the aesthetics of the long-distance views in the neighborhood, but it could be a very small part of the total. Nothing in their study helps clarify what that portion is, nor alters the conclusion of Chestnut and Dennis (1997) that it may be very small.

[30] I estimated what Beron *et al.* found for annual values for a 20% change in ozone, TSP and visual range, respectively, using the coefficients for 1995 in Beron *et al.*'s Table 4 and the 1995 average levels of those variables (from Table 2). My rough calculations closely matched what the authors reported to be the value for a 20% change in visibility (in their Table 7), giving me confidence that my rough estimates are reasonable.

[31] Beron *et al.* (2001), p. 330.

Appendix A offers additional suggestions for ways to explore how to interpret the results in Beron *et al.* (2001).

In conclusion, the Draft ISA appropriately treats the Beron *et al*. study with caution:

> "To the extent the people simultaneously use what they see regarding VAQ as an indicator of the overall air quality including potential health risks, then including all the measures [of air quality characteristics] in the equation is not necessarily sufficient to isolate one effect from the other."[32]

My review of the paper and discussion of its results is intended to more fully explain and further support the caution expressed in the Draft ISA, despite the diametrically opposite conclusions of the paper's authors. My review suggests that even stronger expressions of concern are warranted than the Draft ISA has mustered.

3B. Existence of a New Contingent Valuation Study for Visibility: Smith *et al.* (2005, 2006)

The Draft ISA does not directly review any of the contingent valuation studies of urban visibility, the most recent of which was first released in draft form in 1991 (the final report being McClelland *et al.*, 1993). The contingent valuation method (CVM) was still new at the time the urban visibility studies were being done. CVM techniques have advanced in many ways since then, but only one new major CVM study of visibility values has been performed since then, which is Smith *et al.* (2005, 2006).

Smith *et al.* (2005, 2006) is a CVM study of willingness to pay for regional haze reduction in scenic rural areas, and thus may not be immediately recognized as relevant to urban visibility too. However, this study included a large number of follow-up questions that shed light on how reliably one can interpret the results as a willingness to pay for the aesthetic impacts of rural visibility. It also examined the sensitivity of CVM-based value estimates to alternative survey instrument designs. Several of the insights from the Smith *et al.* study are relevant to the interpretation of the older urban visibility study results.

- Smith *et al.* demonstrate that direct reminders to respondents to consider their household budgets result in a large and statistically significant reduction in the stated values for visibility. The values derived from CVM questionnaires that incorporate budget reminders are only one-sixth of the values derived when the budget reminders were not incorporated.[33] The urban visibility valuation studies lack budget reminders.

[32] Draft ISA, p. 9-83.
[33] Smith *et al.*, (2005), p.1778.

- The new study used follow-up questions to explore the extent to which respondents had intertwined values for perceived health risk reductions in their responses to questions about changes in the aesthetic conditions associated with haze. This concern was highlighted by Chestnut and Dennis (1997) in their review of the urban studies, and Smith *et al.* find a comparably large degree of overstatement due to inclusion of health-related values in respondents' estimates. Respondents reported on average that 78% to 85% of their stated value was actually due to health effects concerns rather than aesthetics except when the questionnaire was designed to elicit a separate willingness to pay for both health and visibility.[34]
- Although the Smith *et al.* study used a scenic vista from a national park, follow-up questions revealed that the respondents were also incorporating their values for comparable improvements that they assumed would also be occurring in the area where they live, and which they would experience in their day to day lives. 35% to 65% of respondents (depending on the questionnaire version) said that *more than half* of their stated value for the parkland visibility change was actually for visibility changes "in urban and suburban areas."[35] In other words, benefits estimates that EPA characterizes as being only the recreational and preservation portion of values from regional haze reduction are based on willingness-to-pay estimates that probably also include residential use values. This means that the summation of benefits for changes in urban and recreational/preservation visibility aesthetics is probably double-counting.

The results in the new visibility valuation study by Smith *et al.* (2005, 2006) provide the first new information since 1991 about how to interpret the results of contingent valuation studies of visibility. The Draft ISA should recognize this new study and summarize its key findings as part of Section 9.3.5.

4. Conclusions

The most important new information provided in these comments is strong evidence that VAQ preference studies do not provide a reliable, enduring or credible estimate of a median level of acceptable visibility conditions. This is based on recent research conducted by CRA International on the sensitivity of the median acceptability cutoff to

[34] Smith *et al.* (2005), p. 1777. The range reflects the averages for different questionnaire designs.
[35] Smith *et al.* (2006), pp. 84-85. This is corroborated by other follow-up questions. One question asked was "*Which one of these statements best describes the types of areas in the Eastern United States where you thought that visibility would improve in the way described by the photographs, when you were answering the questions about cost of living increases*?" Answers were: 15% "in national parks only," 37% "in all rural areas with scenic vistas," and 48% "in some urban and suburban areas as well as rural areas with scenic vistas." Similarly, 86% of respondents indicated the following statement was true of themselves: "*When I answered about how large a cost-of-living increase I would be willing to bear to reduce haze in areas of great scenic beauty, I was consciously thinking that reducing haze in those areas probably means reducing haze over a much greater region as well.*" (p. 81).

changes in the range of VAQ shown in such studies. As no absolute standard appears to exist in people's minds, these surveys do not offer a scientifically credible method for identifying a visibility standard. VAQ preference studies therefore should not be used in setting any visibility standards. The Draft ISA should be revised to include a discussion of this concern, given the new information brought forth in these comments.

These comments have also provided information on a new contingent valuation study that has relevance to the interpretation of other urban visibility valuation studies. Smith *et al.* (2005, 2006) provide the first new information since 1991 about how to interpret the results of contingent valuation studies of visibility. The Draft ISA should recognize this new study and summarize its key findings. Section 9.3.5 should be expanded to include a new section referencing and describing the key results of Smith *et al.* (2005, 2006). This section should include discussion of relevant findings for interpreting the other past urban visibility contingent valuation studies, such as:

- Further evidence on the degree to which health values are embedded in responses to questions about visibility changes.
- The significant change in stated willingness to pay when budget reminders are included in the survey design.
- The evidence that there may be double-counting when residential use studies and recreational/preservation values studies are used to separately estimate total benefits from policies that affect VAQ.

The Draft ISA appropriately treats the Beron *et al*. (2001) study with caution. These comments have provided a detailed discussion of how the results in that paper can be interpreted differently than its authors conclude. This discussion more fully explains the basis for the caution expressed in the Draft ISA. Stronger expressions of concerns are in order.

These comments have also identified a number of corrections that should be made to the Draft ISA. These are:

- The Draft ISA incorrectly summarizes the range of median preferences for an acceptable level of visibility degradation as 19 to 25 dv (or 30 to 55 km). This is not correct because of errors in its summary of results from the Abt Associates (2001) study. The correct range should be stated as 19 to 29 dv (or 21 to 55 km). This error should be corrected on p. 9-4 and p. 9-78.
- The Draft ISA should not characterize the results of preference studies as having a general degree of consistency (p. 9-4, lines 32-33 and p. 9-78, lines 3-4) and instead offer some information on the wide range of $PM_{2.5}$ concentrations that is consistent with a range of 19 dv to 29 dv.
- Errors in how the Draft ISA summarizes the results in Abt Associates (2001) in Table 9-2 (p. 9-77) also should be corrected:

- The entry for "Mean dV found 'acceptable'" under "Washington DC (pilot)" is listed as "~20 dV (range 20-25)" but the correct statement should be "~29 dV (range 27 -31)".

- The entry for "Health issue directions" under "Washington DC (pilot)" says "Health never mentioned, 'Focus only on visibility'." The correct entry in this cell of Table 9-2 should be " 'Harmful to people' mentioned, then 'Focus only on visibility'."

- The Draft ISA provides a misleading synopsis of the nature of health value confounding in the Abt Associates study at p. 9-81, lines 14-19. This paragraph of the Draft ISA should be replaced with a paragraph that notes that the follow-up questions used in that study identified problems on the part of at least a portion of the respondents in separating health from aesthetics in their judgments.

References

Abt Associates, Incorporated. 2001. *Assessing Public Opinions on Visibility Impairment Due to Air Pollution: Summary Report*, work funded through US Environmental Protection Agency, Contract No. 68-D-98-001, Work Assignment 3-53 (January).

Beron, K., J. Murdoch, and M. Thayer. 2001. "The Benefits of Visibility Improvement: New Evidence from the Los Angeles Metropolitan Area," *Journal of Real Estate Finance and Economics*, Vol. 22, No. 2/3, pp. 319-337.

BBC Research and Consulting. 2003. *Phoenix Area Visibility Survey*, Final Report prepared for Arizona Department of Environmental Quality, Denver Colorado (February 13).

Chestnut, L. G. and R. L. Dennis. 1997. "Economic Benefits of Improvements in Visibility: Acid Rain Provisions of the 1990 Clean Air Act Amendments," *Journal of the Air & Waste Management Association*, Vol. 47, pp. 395-402.

Ely, D. W., J. T. Leary, *et al.* 1991. "The Establishment of the Denver Visibility Standard," paper 91-48.4, presented at the 84th Annual Meeting of the Air & Waste Management Association, Vancouver, British Columbia (June 16-21).

McClelland, G., W. Schulze, *et al*. 1993. *Valuing Eastern Visibility: A Field Test of the Contingent Valuation Method*, report prepared by University of Colorado under USEPA Cooperative Agreement #CR-815183. Washington, DC (September).

Pryor, S. C. 1996. "Assessing Public Perception of Visibility for Standard Setting Exercises," *Atmospheric Environment*, Vol. 30, No. 15, pp. 2705-2716.

Smith, A. E., M. A. Kemp, *et al.* 2005. "Methods and Results from a New Survey of Values for Eastern Regional Haze Improvements," *Journal of the Air & Waste Management Association*, Vol. 55, pp. 1767-1779.

Smith, A. E., M. A. Kemp *et al.* 2006. *Valuing Regional Haze Changes: The Sensitivity of Contingent Valuation Results to Questionnaire Format*, Final Technical Report (1011852) prepared for Electric Power Research Institute, Palo Alto, California (March).

Trijonis, J., M. Thayer, *et al.* 1985. *Air Quality Benefit Analysis for Los Angeles and San Francisco Based on Housing Values and Visibility*," Final Report prepared for Research Division of the California Air Resources Board, ABA Contract No. A2-088-32, Sacramento, California (January).

Appendix A. Additional Suggestions for Further Assessment of Beron *et al.* (2001)

When a new study offers dramatically different results from a predecessor study, it is useful to try to understand what the root causes are for the large changes. A couple of suggestions are offered below for attempting to compare Beron *et al.* (2001) to its predecessor, Trijonis *et al.* (1985), with the specific focus on whether the newer study has better separated health effects from aesthetic effects.

First, one might ask whether the large increase in value between the Trijonis *et al.* study and the Beron *et al.* study might reflect an increasing concern among the general public that air pollution poses health risks. The earlier study reflects residents' values in 1978-1979, which predates most of the literature that reports to find that air pollution may pose risks of premature death. The later study reflects values over the period 1980-1995. Public awareness and concerns with air pollution rose significantly over this period (even while air pollution was falling), and to the extent that both studies' visibility value reflects those health effects concerns, one would expect those values to increase more rapidly than any income effects would indicate.

Second, it would be interesting to apply the current statistical techniques to the data in the earlier study (and/or to apply the older techniques to the current study). One important difference from the earlier study is the incorporation of TSP and ozone as explanatory variables in addition to visibility. If the same air pollution data were to be developed for the 1978-9 period, then the Trijonis *et al.* could be replicated, and then checked to see if its findings would change as a result of including TSP and ozone. Of particular interest would be whether the values attributable to the air pollutants had increased in a comparable manner to the changes in the visibility values. A finding that the TSP and ozone effects were generally similar between the earlier and later study, while the visibility component increased substantially, would indicate that perceptions of health benefits are embedded more in the visibility-related value and not in the estimated values the for the air pollution variables.