Peer Review for RTI Report, "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge?

United States
Environmental Protection
Agency

Peer Review for RTI Report, "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge?

Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

Prepared for EPA by
RTI International
EPA Contract No. EP-C-16-021
Work Assignment No. 1-10
and
ICF International
EPA Contract No. EP-C-16-020
Work Assignment No. 1-16

NOTICE

This technical report does not necessarily represent final EPA decisions or positions. It is intended to present technical analysis of issues using data that are currently available. The purpose in the release of such reports is to facilitate the exchange of technical information and to inform the public of technical developments.

United States Environmental Protection Agency

EPA-420-R-18-017
July 2018

July 2018

MEMORANDUM

SUBJECT: Peer Review for RTI report, "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge?"

In August 2017, EPA contracted with ICF to conduct a peer review of a study conducted by RTI International (RTI). The draft study, titled "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge?" derived estimates of consumer willingness to pay (WTP) for vehicle attributes from existing studies of vehicle demand and WTP for attributes.

The three peer reviewers selected by ICF were Drs. David Brownstone of the University of California at Irvine, David Bunch of the University of California at Davis, and George Parsons of the University of Delaware. EPA would like to extend its appreciation to all three reviewers for their efforts in evaluating this survey. The three reviewers brought useful and distinctive views in response to the charge questions.

The first section of this document contains the final RTI report responding to the peer reviewers' comments. It begins with responses to the summary of comments provided by ICF, and then presents responses to the detailed comments from each peer reviewer. The second section provides the peer review report conducted by ICF. It documents the peer review process, provides both a summary of the peer review comments and the detailed responses to charge questions, the peer reviewers' resumes, letters from each reviewer explaining any real or perceived conflicts of interest, and summary notes from a mid-review meeting.

CONTENTS

I. Authors' Responses to Peer Review of "Consumer Willingness to Pay for Vehicle Attributes"—Final Report (June 2018), conducted by RTI International.

1. Summary Comments
2. Detailed Comments
3. References
4. References

II. Peer Review of "Consumer Willingness to Pay for Vehicle Attributes" – Final Report, conducted by ICF.

1. Introduction
2. Peer Review Process
3. Responses to Charge Questions

Appendices

A. Resumes of Selected Peer Reviewers
B. Conflicts of Interest Declarations
C. Peer Reviewer Mid-review Meeting Note

RTI INTERNATIONAL | Memorandum

To: Gloria Helfand, U.S. Environmental Protection Agency

From: David Greene, Anushah Hossain, Julia Hofmann, and Robert Beach

Date: June 28, 2018

Subject: Authors Responses to Peer Review of "Consumer Willingness to Pay for Vehicle Attributes"—Final Report (June 2018)

We thank the reviewers for their conscientious expert review of our report. Their queries and recommendations have enabled us to improve the report by correcting oversights, adding new material, and clarifying subjects that were inadequately explained. This memorandum explains our responses to the reviewers' comments. It is comprised of two parts. In the first part we respond to the summary of the reviewers' comments on each of the four prescribed questions and describe changes we have made to the document as a result. In the second, we address the comments by reviewer.

1 Summary Comments

1.1 Is the Choice of Publications Described Sufficiently?

We appreciate the reviewers' general assessment that the choice of publications was described sufficiently and two reviewers' observations that the choice of studies appeared thorough.

In response to Dr. Bunch's suggestions to provide more detail on how we determined the set of studies included in our main sample, we have added explanation to the report that we believe better explains our rationale for limiting our scope to U.S. studies and dates between 1995 and 2015. Our report is intended to inform the EPA's current policy making for the United States. While there are undoubtedly important insights to be found in studies of consumer demand for vehicles in other countries, adding willingness to pay (WTP) estimates from other countries would add additional dimensions of uncertainty due to different tax regimes, different geographies, and different cultures and preferences. Given limited time and resources, we elected to focus on studies based on U.S. data. We limited our scope to studies published in the 1995–2015 period to reduce the impact of changing preferences over time. Limiting our scope to more recent studies also ensures greater uniformity and sophistication of inferential methodology. We limited our selection of papers to those published in peer-reviewed journals or in the gray literature published by academic or research institutions. Our intent was to include any studies that met these criteria and provided sufficient information to calculate WTP values.

We are encouraged by the conclusions of two reviewers' that our search and selection was "thorough." Although Dr. Bunch commented that he is aware of some studies that were excluded, he did not provide any guidance on what he thought was missing or mention any specific studies. Therefore, we did not change our selection of papers in response to the peer review comments.

RTI International is a registered trademark and a trade name of Research Triangle Institute.

All three reviewers offered suggestions for improving the descriptions of the analyses we included in our survey. They requested additional information on sample sizes, the populations on which inferences were based (e.g., new car buyers, all car buyers, all households), and a frequency distribution of studies for the top 5–10 journals in which studies are most frequently published. In response, we have added a table listing all 52 studies with these and other descriptive factors included. We agree that presenting such information will help readers understand the nature of the studies we reviewed. In particular, presenting the sample sizes, sample types, and populations included will help readers understand how varied the sources of data are. We believe this will help readers understand our decision not to weight WTP estimates by sample size because sample sizes are, in general, not comparable across studies. One could say that the sample size of the Berry, Levinsohn, and Pakes (1995) paper is 20 by counting each year from 1971 to 1990. On the other hand, in each year there are approximately 110 makes and models included, making the sample size around 2,200 instead of 20. However, all automobile sales (aggregated by make and model) are included in the study, making the sample qualitatively different from, for example, the study by Busse, Knittel, and Zettelmeyer (2013) that had a sample size of just under 2 million individual automobile purchase transactions that occurred between 1999 and 2008. The question, "Which sample is larger?", does not have a simple answer because one study is based on 20 years of aggregate sales data comprising the total population of vehicle purchases (approximately 15 million per year) while the other is based on a very large sample of individual purchase transactions over a 10-year period. We agree that the reviewers' recommendation to provide additional information of this nature will help readers better appreciate one reason why WTP estimates are as variable as they are.

We thank Dr. Parsons for calling attention to our attempt to improve the quality of our analysis by contacting every study's lead author, providing the author with a spreadsheet containing our WTP calculations, and requesting comments or corrections. We adopted this procedure in addition to querying authors when we were unsure of units of measure, definitions, or other key pieces of information. To our knowledge, this has not been done before in literature reviews such as ours. It was motivated by the complexity of WTP calculations in many cases, as well as the highly varied model formulations we encountered. We know that the process was effective in several cases, resulting in corrections or modifications of our calculations. We are ultimately responsible for the quality of our analysis, but we are grateful to the authors who took the time and effort to respond to our requests, which helped us improve the study.

1.2 Are the Methods Described Sufficiently?

We are glad to see that the reviewers concluded that our methods were described in sufficient detail. We also thank them for their suggested improvements. Based on these suggestions we modified our text in the following ways. We have attempted to clarify the purpose and goals of our study. The goal of our study was to establish what information the existing literature provided concerning consumers' WTP for vehicle attributes to inform the EPA's decision making. Although it was not the goal of the phase of our study described in this report to conduct a meta-analysis of those estimates, we intended from the beginning to conduct a focused meta-analysis in a follow-on study (as conducted in Greene et al., 2018). For that

reason, we followed well-known procedures for meta-analysis of WTP estimates, essentially identical to methods described by van Houtven (2008). To make this clearer, we have added to our report the description of the steps recommended by van Houtven that we followed in selecting studies for inclusion, designing and creating our database and calculating WTP estimates.

There are two areas in which our approach differs somewhat from van Houtven's paradigm. We limited our judgment about the quality of studies to the requirements that either the study be published in a peer-reviewed journal or that it be a gray literature study of similar quality based on authorship and publishing organization. Selecting studies based on the analysts' judgment concerning the quality of the research could improve the quality of the WTP estimates and perhaps reduce their variability. On the other hand, it will also introduce the analysts' confirmation biases. Because the primary goal of this report is to describe the evidence on WTP for vehicle attributes in the existing literature we chose to err on the side of avoiding confirmation bias while ensuring that the studies included met the standard of professional journal peer review. We acknowledge that by including some studies from the gray literature we have introduced our own judgment about whether they meet that standard. However, including studies from the gray literature is also frequently recommended in designing meta-analyses to mitigate "publication bias" (van Houtven, 2008, p. 904). We are encouraged by the fact that none of the reviewers suggested that any of the gray literature studies we included should not have been included.

The second area in which our method differs slightly from van Houtven's paradigm is that we do not use sample size as a criterion for inclusion (nor did we use it in our later meta-analysis to weight WTP estimates). Instead, we used a measure of the precision of the estimated attribute coefficient (or derivative); that method is addressed at length in Section 1.3 of this response to comments. Our reasoning here is due to the different meanings (and enormous variability) of sample sizes across studies. As discussed above in Section 1.1 of this response, aggregate market sales sample sizes are not directly comparable to individual transactions data sample sizes, which differ from revealed preference survey sample sizes, which are also qualitatively different from the sample sizes of designed stated preference experiments. The use of sample size as a primary measure of the quality of a metric derives from the application of meta-analysis to designed experiments, such as clinical trials. The use of sample size to weight, say, outcome metrics in such cases is quite reasonable. When methods, models, and data vary as widely as they do in our case, sample size is an inconsistent indicator of the reliability of statistical inferences.

As recommended by Dr. Parsons we have also expanded the discussion of methods of estimating WTP to explicitly include nested multinomial logit (NMNL) models. In all random utility models, we use the ratio of the derivative of the indirect utility function with respect to the attribute in question to the derivative of vehicle price to estimate WTP. In the case of NMNL models, we take the derivatives at the nesting level at which the attribute in question enters the model. It is certainly possible and, in fact, generally the case that there are multiple WTP estimates for NMNL models whenever the same attribute is included in many nests.

1.3 Are the Methods and Procedures Technically Appropriate?

We are pleased that the reviewers found our methods and procedures to be technically appropriate. We appreciate their suggestions for improvement and we would have adopted many of them were it not for the absence of relevant statistics in the published literature.

We and most others in the peer-reviewed literature estimate WTP as a ratio of estimated coefficients (more generally, the ratio of derivatives of the utility function with respect to an attribute and to the vehicle's price). Because coefficients are estimated with uncertainty, such ratios are ratios of random variables. The mean of the ratio of two random variables is not, in general, the ratio of their means, raising methodological issues. We acknowledge this issue in the report. A few of the authors of papers we reviewed recommended we use the Delta method to estimate the mean of the ratio. As we also note in the report, the benefit of the Delta method for estimating the ratio of two random variables derives from knowing their covariance. However, published articles and reports almost never provide the estimated variance-covariance matrix for coefficient estimates. Using the Delta method without including the covariance introduces an unknown bias. Thus, we have a dilemma: estimate the mean of the ratio with an unknown bias by using a simple ratio of means or introduce a different unknown bias by using the Delta method without the estimated covariance.

Our solution is to estimate WTP *conditional* on the central tendency estimate of the price derivative. While this solution is less than ideal, it has the advantage of having a meaningful interpretation and of being both feasible and an accurate description of the WTP metric. It also allows us to measure the uncertainty of this metric using the standard error of the attribute derivative. Although we acknowledge the concerns expressed by some of the authors of papers included in this study and reviewers, we are aware of no better, feasible alternative.

In our concluding section we recommend that authors routinely calculate and report WTP metrics in their published papers using the most accurate method, and some are already doing so. Because the authors have access to all the relevant statistics, this should result in improved central tendency estimates of WTP. The reviewers' comments have led us to revise the description and, we hope, successfully clarify our method and reasoning.

We thank Dr. Bunch for directing us to the unpublished paper by Carson and Czajkowski (C&C) (2013). The paper is focused on the problem of taking an expected value of a ratio whose denominator has some finite probability density at zero. This is potentially a problem for estimating WTP from random utility models of vehicle choice because the derivative of price (which is often the price coefficient alone) is estimated with error, the distribution of which is most often assumed to be asymptotically normal. C&C's proposed solution is to estimate models in a way that constrains the estimated coefficients to have no probability density at zero (e.g., assume lognormally distributed coefficient estimates, use truncated normal distributions). Unfortunately, this is not feasible for a study such as ours, which relies on estimates reported in published literature by authors. Moreover, the probability distribution of an estimated coefficient is typically a premise rather than an inference. In reality, there is no realistic expectation that

the derivative of utility with respect to vehicle price is, in fact, zero. We believe that (with very rare exceptions) the existence of a finite probability density at zero for the price derivative is an artifact of the estimation methods. Indeed, the solution proposed by C&C is to assume an alternative error distribution that has no probability density at zero. Our interpretation of WTP estimates as conditional on the central tendency of the price derivative avoids this problem while also providing a relevant measure of WTP. Although our metric is not perfect, it is computable, meaningful and, as the reviewers note, the most commonly used metric in the peer-reviewed literature. We have attempted to clarify our explanation.

Dr. Brownstone's suggestion that models be estimated in WTP space solves both parts of the ratio of random variables problem. It was an oversight on our part not to mention this more prominently, especially since one of the studies in our database (Train and Weeks, 2005) does estimate its coefficients in WTP space. We have added a brief discussion of this in our revised report. Unfortunately, Train and Weeks (2005) is the only study among the 52 in our main sample that used that method.

Dr. Bunch recommends that we use the median as a measure of central tendency. We do this only in the case of random coefficient models for coefficients assumed to be lognormally distributed. When distributions are skewed, like the lognormal, the median is generally considered a better measure of central tendency than the mean. In many cases, the distributions of the full set of WTP estimates we have derived from all the studies are skewed. For this reason, in our summary tables of WTP estimates we report not only means and standard deviations but also skewness, medians, and the interquartile ranges.

We thank Dr. Brownstone for his recommendations for estimating standard errors of WTP estimates. However, given the interpretation of our WTP estimates as conditional on the point estimate of the price derivative, we have decided not to use this method. We agree that our method underestimates the standard errors of the true mean of the ratio of random variables and we have attempted to emphasize this point in our revised text. Given the definition of our WTP metric, we believe the standard errors of the attribute derivatives are a more appropriate measure of uncertainty.

Dr. Parsons recommends the use of simulation methods to estimate unbiased means. We concur with his reasoning but note that the information necessary to carry out such simulations is generally only available to the authors of a study. In our concluding section we have revised our text to make this a key recommendation to authors.

Dr. Parsons recommended that we should focus on data from stated preference studies. His points about the problems with sample design and identification of the demand function in revealed preference studies are well taken. However, we believe that stated and revealed preference studies have different strengths and weaknesses. Revealed preference studies generally suffer from less-than-ideal sample designs. Omitted variables and errors in variables are also likely to be problems. Stated preference studies are subject to "hypothetical bias" and although there are procedures to counter it, one can never be certain that it has been reduced to a negligible level (Loomis, 2014). The degree of hypothetical bias also tends to vary depending on how questions are phrased, as well as varying from one attribute to another. Because our study's primary goal is to describe what the published literature says about WTP for vehicle

attributes, we believe it is best to present estimates from both sources. For fuel cost and performance, we present summary statistics separately for studies based on stated versus revealed preference survey data, and also for market sales data. In our follow-on meta-analysis (Greene et al., 2018), we include separate assessments of WTP estimates from revealed and stated preference studies, thereby enabling the EPA and others to apply their own judgments about which type of study is most reliable.

1.4 Are the Results Presented Appropriately?

"The reviewers all agreed that the results were presented appropriately and are generally expected given the common findings in the literature." We are glad to see that the reviewers provide general confirmation of our approach to presenting results and thank them for sharing their concerns and suggestions for improvements. In particular, as discussed above, Dr. Bunch's recommendation that a meta-analysis be conducted is something we planned from the beginning as a separate follow-on study. That research has been completed and is currently under consideration by a peer-reviewed journal. Now that the study exists, we have mentioned it in our report and have added a reference to the paper that is currently in the process of peer review.

Dr. Bunch's recommendation that EPA calculate marginal rates of substitution between attributes such as fuel economy and performance is interesting, and worth consideration. However, it is outside the scope of our study. We note that the same methodological issues raised in Section 1.3 would also arise here because marginal rates of substitution would also generally be the ratio of two random variables. In addition, marginal rates of substitution, defined as the ratios of marginal utilities of two attributes, can be computed from our results as the ratio of two WTP estimates. Proper interpretation of such WTP estimates would still be an issue, however.

We agree that providing more background discussion concerning the economic theory of WTP derived from indirect utility functions would be helpful to some readers. We have therefore added this to the theoretical discussion. Calculating WTP from logsums (e.g., from changes in consumers' surplus) is an interesting idea because it can estimate WTP for discrete rather than marginal changes in attributes. We agree this is worth considering as standard practice for authors in future research. It is also related to Dr. Parsons' recommendation to carry out simulations analyses, discussed in Section 1.5.

As noted above, Dr. Brownstone's recommendation to estimate models in WTP space solves the problem of characterizing the ratio of two random variables. It should be considered by future authors as potentially a superior approach to estimation in marginal utility space whenever the objective of the research is estimating WTP.

Dr. Parsons' observations about WTP for vehicle class are well taken. We have improved our explanation of why we calculated but do not present measures of WTP for vehicle class. Our reasons are the same as his. WTP estimates for vehicle class depend on which class is chosen as the default. Not only does this vary from one study to another but the numbers of classes and definitions of classes also vary so much

that comparable WTP estimates are almost non-existent. For fuel type, we note that all the WTP estimates are relative to a conventional gasoline vehicle.

1.5 Additional Comments

We appreciate Dr. Bunch's recommendation that we carry out a meta-analysis of the WTP estimates. As noted in Section 1.2, we intended from the outset to perform a meta-analysis of at least some of the WTP estimates in the second phase or our research and we structured our literature search, data collection, and estimation methods with that in mind, following van Houtven (2008). We have added material to the report to make this clear. However, it was not our plan to include the meta-analysis in this report. The goal of this report is to describe our methods in detail and provide a description of the resulting WTP estimates. After finishing the report, we conducted a meta-analysis focused mainly but not exclusively on WTP for fuel economy and performance, the two measures most frequently represented in the literature (after vehicle price). Serendipitously, the trade-off between fuel economy and performance may be the most relevant with respect to greenhouse gas standards. Again, the weights we used in the meta-analysis were not sample size, for reasons explained in Section 1.2.

Dr. Brownstone is correct in noting that some of the studies we included were not carried out with the intent of estimating WTP for vehicle attributes but rather for other purposes, such as prediction or policy evaluation. In other studies, the authors may (or may not) assert that certain variables (e.g., vehicle weight or dimensions) are included as proxies for other excluded attributes. Again, because this report is intended to describe existing evidence on the WTP for vehicle attributes, and because our intent was to avoid introducing confirmation bias, we did not attempt to judge which estimates were "good" estimates of WTP and which were not. However, we did exclude estimates from models that authors clearly indicated were incorrect or invalid.

Although the alternative methods of policy evaluation proposed by Dr. Brownstone are also useful metrics, the statement of work for this study was specifically focused on describing the evidence from the extant literature concerning the marginal WTP for vehicle attributes. The alternative methods recommended also have strengths and weaknesses and ultimately also depend on the validity of a particular model and its estimated parameters. Again, we agree that estimating discrete choice models in WTP space removes the problems associated with the ratios of random variables but researchers have generally not used that approach and it was not possible for us to impose it, *ex post*.

We thank Dr. Brownstone for pointing out two mis-statements. First, while A/C has become standard in nearly all new vehicles, that does not mean that it is an inconsequential attribute. The near ubiquity of A/C may make it more difficult to estimate its value but that is not a reflection of its value. Second, the statement about the relative values of WTP for fuel cost reduction from stated and revealed preference sources should be comparing both to market-sales-based estimates. We have corrected these errors.

Dr. Parsons' observations about the complexity of the report are understandable. It is quite a challenge to describe in a simple way the WTP for all vehicle attributes. With respect to focusing on a few key

metrics, we have taken this approach in the follow-on meta-analysis, in which we focus primarily on fuel cost and performance. We hope that study is both less complex and more clearly presents key WTP estimates relevant to regulatory policy.

We think Dr. Parsons' suggestion to present WTP estimates for relevant consumer groups is an interesting approach that is likely to produce policy-relevant information. While it is beyond the scope of the current study, we will bear it in mind as a potentially valuable avenue for further research.

We considered using only the model from each paper that was identified by the authors as the preferred model. Instead, we included all models that the authors did not identify as clearly incorrect on methodological or statistical grounds. However, we recorded which model was identified by the authors as the preferred model as a potentially useful variable in a meta-analysis. In any case, we believe that authors who presented more than one plausible, defensible model formulation have done a service in advancing our understanding of why WTP estimates vary so widely across studies. In our meta-analysis these studies provided greater diversity of methods and assumptions with which to identify causes of variability across studies.

We agree with Dr. Parsons that including multiple models from a single paper can give such studies greater influence in determining measure of central tendency or variability. The same applies to multiple papers using the same data source. In our meta-analysis work we tested weighting by the inverse of the number of studies from the same paper or the number of studies using the same database. In the end, we settled on weighting by our measure of the uncertainty of each estimate because it was more consistent with the very strong finding of random effects across studies. In the context of meta-analysis by the DerSimonian & Laird method we used, the term "random effects" implies that the various WTP estimates are in fact measuring different "true" parameters. This is very consistent with the fact that the studies, in general, come from different periods and include different populations. Furthermore, differences in model formulations and estimation methods almost certainly introduce model-specific biases.

Dr. Parsons' suggestion to conduct a structured set of simulations using a set of carefully chosen models to estimate the welfare effects of various policies is interesting and worthy of consideration. Such a method raises its own methodological issues, however, and is outside the scope of our study. Again, this is a suggestion worth considering for future research.

2 Detailed Comments

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	With the above discussion as background, I would first comment that, although the authors provide documentation on their process for identifying and collecting studies, I have concluded that there is not enough detail on the reasons for why some of the selection criteria were chosen. I review again here that, based on my (admittedly very quick) review of some of the meta-analysis literature, I think all of the steps performed in this work could have benefited from more structure and rigor of the type I saw described in that literature.	This comment is largely addressed in the first paragraph of Section 1.1. The procedures we followed in selecting studies, collecting data about them and developing WTP estimates are the same procedures followed when preparing to perform meta-analyses. In the current report we show how the steps we took closely correspond to the paradigm recommended by van Houtven (2008) (see page 1-2 of the report).
	In terms of specifics: The studies are limited to those involving US-based populations. The argument here is (presumably) that, because the policies are for the US market, there would be no value in including studies from, e.g., Europe. On the surface, this seems understandable and perhaps even reasonable. However, given the well-known concerns about the variability of WTP measures and the potential impact of sample sizes, research methodologies, etc., the existence of such studies could provide additional statistical power for evaluating these concerns, which I think are paramount. On the other hand (and in fairness to the authors), I suppose it could be argued that the extent of these issues might not have been fully appreciated ex ante, and that the extent of these concerns might not have been fully evident except as an outcome of this work. Having said this, this actually reinforces the importance of treating this aspect of the design process in a more rigorous way.	Our approach to selecting studies and estimates within studies appears to be the main concern of this comment. We deliberately cast a wide net when including studies. However, we were selective in that we focused primarily on peer reviewed studies but included studies from the gray literature that were, in our opinion, of similar quality. We have expanded the discussion of this subject to respond to this comment. We did not rank or weight studies based on our perception of quality. We did include a measure of the rank of the journal in which the study appeared, however, as a factor for use in the subsequent meta-analysis. We did not selectively choose WTP estimates within studies based on our subjective judgment. We wanted a large sample of WTP estimates and we wanted to explore what the literature had to offer. We also wanted to avoid introducing confirmation bias in a way that could influence our results. We have made changes to the relevant sections of our report to better explain our goals and methods. We could have eliminated estimates that were statistically insignificant at a certain level or had a sign that conflicted with prior expectations. We could have tried to discern which estimates the authors believed reflected a true WTP versus those that were introduced in an attempt to represent variables unavailable to the

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
		author. We could also have tried to determine whether studies were or were not attempting to realistically estimate consumers' preferences or were aimed at prediction or some other goal. We considered all those possible approaches and decided that, for a first comprehensive review of the literature, it was best to err on the side of inclusivity and objectivity, to describe what evidence exists and to keep the potential for confirmation bias to a very low level. We have made a number of revisions in response to this comment and hope we have done a better job of explaining this approach in our revised paper. Our reasoning for using only U.S. studies is explained in the first paragraph of Section 1.1 of this memorandum as well as Section 1 of the report. We attempted to achieve a balance between the number of studies included and the fact that consumers' preferences may differ geographically.
	Similarly, a decision was made to use a cutoff of 1995. The rationale provided is not compelling, particularly if this cutoff were to eliminate studies that might have disproportionate value due to superior sample size, potential study quality, etc., given the well-known concern about potential variation in WTP estimates. For example, there are some relatively important studies that appeared in the small number of years prior to 1995 that could have been important to include.	Our reasoning for selecting the cutoff date of 1995 is explained in the first paragraph of Section 1.1 of this memorandum as well as Section 1 of the report. We attempted to achieve a balance between the number of studies included and the fact that consumers' preferences have likely changed over time.
	More generally, as I reviewed this work it seemed to me that there were studies that I am personally aware off that were probably not included, but under the circumstances it was difficult to sort this out (and I believe it would have been inappropriately time consuming). These studies could have been omitted for various reasons stated by the authors, but there was no efficient way for me to know which studies were eliminated, and why.	As with any literature review and synthesis, there may have been additional studies that met our criteria and ideally should have been included in our survey. However, we are confident that we did a rigorous literature search combining all the methods recommended by van Houtven (2008). We are pleased that Dr. Brownstone commends our work in this area: "…– they appear to have done an excellent job finding all of the relevant studies…". Absent specific recommendations concerning which studies we should have included, we did not adjust the set of studies included in the analysis.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	In summary (and to directly address this specific charge question, as stated): I found that the current version of this review makes it rather difficult to evaluate the "quantity and quality of the data used for analysis." This response can also be linked back to some the introductory remarks provided above. The issue of sample size, in particular, has not been addressed in this study. More generally, the authors have collected a lot of measurements here, and it seems that it would be possible to do a more systematic exploration of a variety of important issues that have been identified, e.g., the issue of whether the type of data (e.g., RP, SP, market data, etc.), amount and structure of data (intertemporal and/or geographic variation), and modeling methodology systematically affects these estimates, and, if so, in what way. This could lead to more carefully considered conclusions on ranges for WTP estimates. (See later comments in response to other questions.)	We have added more summary statistics for our main sample in Section 3 of the report. As discussed elsewhere, there are a number of reasons why sample size cannot readily be used to weight across the various types of studies incorporated in this review. As noted in Section 1.2, we have now separately completed a meta-analysis of selected WTP estimates and we now mention that in our report. Presenting the results of that analysis was outside the scope of the work described in this report.
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	The short answer to this question is a lightly qualified "yes." Based on my exploration of the material in Section 4 and Appendix C, I have some potentially serious concerns about how the calculations were performed in this study. I will go into greater technical detail in my response to the next question. However, as I general matter, I would echo my earlier remarks that my review of various literature references has led me to conclude that the design and execution of this study could have benefited from application of the more formal structure that has been developed in the meta-analysis literature in general, and on the topic of WTP in particular. This	As explained in Section 1.1 of this memorandum, we did follow the generally accepted procedures for meta-analysis because we anticipated doing a meta-analysis in a second phase of the study. We now explain this by referencing the steps outlined in van Houtven (2008). This will help readers better understand our methods. The difficult issues of how to estimate central tendency WTP values and to describe their uncertainty is explained in Section 1.3 of this memorandum. We understand that our measure of uncertainty is an underestimate of the uncertainty of the mean of the WTP estimate. However, because the published literature almost never supplies enough information to calculate an unbiased estimate of the mean in the general case where the estimated attribute and price derivatives are correlated, we interpret our estimate as conditional on the central tendency value of

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	was already mentioned in the response to charge question 1 regarding the selection process for studies. In the next charge question, I will go into more detail on the technical question of computing WTP measures. But, more generally, the overall process should pay attention to issues related to computing the ES measures, and how to compute measures of uncertainty both within and across studies. This can culminate in development of more systematic models of these measures that shed light on the nature and source of these variations. Again, although I realize that this specific study clearly intended to "stop short" of performing so-called "meta-analysis regression modeling," I have concluded the procedures for properly performing the initial steps could have greatly benefited from adhering more closely to the meta-analysis paradigm that exists in the literature.	the price derivative. Nonetheless, we find that even using this interpretation and using one rather than two standard deviations of the attribute derivative, uncertainty bounds are large relative to the mean value in many cases. In the report being reviewed here, we focus on the variability of the central tendency measures of WTP across papers and models.
3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	A critical aspect of this study is that the vast majority of WTP measures are computed as ratios of estimated coefficients, and the statistical framework for producing the results is based on the Delta method as discussed in Section 4 and Appendix C.	In general, median coefficient estimates are available only for MXL models whose coefficients are assumed to be log-normally distributed. For MXL model coefficients that were assumed to be log-normally distributed, we did use median values in calculating derivatives. We discuss our rationale for the statistical measures used in the report as well as Section 1.3 of this memorandum.
	Although it seems to be not very well appreciated in much of the social science literature, in the statistics literature there are well-known problems that arise when a statistical measure is computed as the ratio of two normally distributed random variables: the distribution of this ratio is Cauchy, with undefined mean and undefined variance.	We discuss our rationale for the statistical measures we used in the report as well as Section 1.3 of this memorandum.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	In the report, the authors compute WTP measures using a ratio of estimated coefficients from discrete choice models in the vast majority of cases. The coefficient estimates have typically been obtained by maximum likelihood, and are treated as being asymptotically normal with some variance-covariance matrix. In various subfields of the social science literature, researchers have commonly used the Delta method (or a variant) for estimating statistics related to the WTP distribution. Another method includes parametric bootstrapping (Krinsky and Robb, 1986).	We discuss our rationale for the statistical measures we used in the report as well as Section 1.3 of this memorandum.
	However, statistical properties of quantities computed using these methods can be very badly behaved unless the estimated coefficient for the denominator variable (i.e., the price coefficient) is "statistically far" from zero. One result from the literature suggests the following rule of thumb: To use the Delta method, the t-statistic for the denominator coefficient should be above 8.75. (Note that this depends on the combined values of the coefficient estimate and the estimated standard error.) Now, there has been a variety of Monte Carlo studies in the social science literature that suggest the Delta method "works well." However, in the fine print it can be discovered that they frequently generate results using cases where the denominator coefficient has been estimated with a high degree of precision (as suggested by the rule of thumb mentioned above).	We believe our interpretation of the WTP values as conditional on the mean (or median) value of the price derivative is the best choice among available alternatives. We recognize that our measure is generally not the mean value of the WTP estimate. On the other hand, it can be computed from the available information and has a meaningful interpretation. We have amended our text to explain this more clearly.
	Before continuing, one specific recommendation to the authors is that they obtain and study a working paper by Carson and Czajkowski (2013) ["A New Baseline Model for Estimating Willingness to Pay	We thank Dr. Bunch for directing us to the unpublished paper by Carson and Czajkowski (C&C) (2013). The paper is focused on the problem of taking an expected value of a ratio whose denominator has some finite probability density at zero. This is potentially a problem

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	from Discrete Choice Models"] where the technical details mentioned above are discussed in more detail. In particular, they offer a simulation example showing that, when the coefficients estimates have been obtained with only moderate precision, that there are specific problems with the behavior of both the Delta and Krinsky-Robb methods. In particular, the mean and variance statistics should be avoided (as might be expected given that the statistic is, after all, Cauchy distributed). The median is the appropriate measure of central tendency, the 96% confidence interval/quantile range contains negative numbers, and it is skewed. At the same time, the Delta method yields a reasonable mean (even though it shouldn't), and a standard deviation that is not only finite (which is wrong) but also overly small, leading to confidence intervals that are both inappropriately narrow and symmetric (also wrong).	for estimating WTP from random utility models of vehicle choice because the derivative of price (which is often the price coefficient alone) is estimated with error, the distribution of which is most often assumed to be asymptotically normal. C&C's proposed solution is to estimate models in a way that constrains the estimated coefficients to have no probability density at zero (e.g., assume lognormally distributed coefficient estimates, use truncated normal distributions). Unfortunately, this is not feasible for a study such as ours, which relies on estimates reported in published literature by authors. Moreover, the probability distribution of an estimated coefficient is typically a premise rather than an inference. There is no realistic expectation that the derivative of utility with respect to vehicle price is, in fact, zero. We believe that (with very rare exceptions) the existence of a finite probability density at zero for the price derivative is an artifact of the estimation methods. Indeed, the solution proposed by C&C is to assume an alternative error distribution that has no probability density at zero. Our interpretation of WTP estimates as conditional on the central tendency of the price derivative avoids this problem while also providing a relevant measure of WTP. Although our metric is not perfect, it is computable, meaningful and, as the reviewers note, the most commonly used metric in the peer-reviewed literature. We have attempted to clarify our explanation.
	Taking a step back, and repeating from before: Recall that a major outcome of this report is the observation that the computed WTP results seem to "vary widely," and that this raises a variety of concerns. In addition, the authors also note that, simultaneously, the computed estimates of "precision" for WTP measures within a given study are rather high, which seems to be a contradiction. However, note that this outcome would appear to be entirely consistent with the discussion given above regarding what can happen when using the Delta method in this situation. Specifically, Carson and	We believe the correct interpretation of the result that estimates vary widely across studies while the standard errors of attribute derivatives are more precise within studies is that the studies are not measuring the same "true" WTP values. This is because of the various biases created by omitted variables, errors in variables, and model formulations as well as to differences in populations sampled and time periods. We have edited the report text to make this clearer. We disagree with the assertion that our procedure for estimating WTP values is "ill-advised". As we explain in Section 1.3 of this memorandum, our approach computes a meaningful and interpretable metric. We have modified the text of the report to provide a better explanation of the method and its strengths and weaknesses.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	Czajkowski (2013) explains in detail why the procedure followed by the authors is ill-advised.	
	Have gone over this, I would add that I don't have the wherewithal to give the authors step-by-step instructions on exactly how to do this instead. I want to be clear that I also understand appreciate that the authors are facing constraints in being forced to use only the information that is available from existing published studies. However, I do have some ideas about possible directions. At the very least, it is clear that the measures of central tendency should probably be the median and not the mean to the degree this is possible, and some alternative approach to computing a confidence interval (or perhaps, e.g., an interquartile range) should be used for measuring uncertainty (i.e., not the standard deviation). The material in Section III of Carson and Czajkowski (2013) would seem to provide some possible avenues for doing this. Having said this, we must acknowledge that almost any procedure developed on this basis could run into the same difficulties the authors faced in this study, i.e., that certain key statistics (such as covariance estimates) might not be available.	We do use the median price coefficient value in MXL models, which is generally the only case where the median can be calculated from available information. However, many MXL models assume fixed (nonrandom) price coefficients.
	More generally, the authors should take a more careful look at the literature on WTP statistical measures, and also at actual meta-analyses of WTPs that appear in the literature. [However, with regard to the statistics, my view is that the Carson and Czajkowski (2013) paper should be the touchstone: There are some published results in the transportation literature that one should be careful about when considering, e.g., there is a paper by Gatta et al. (2015) in Transportation Research A on	Thank you for raising these issues. We now cite the paper by Daly et al. 2012 even though we cannot use the recommended method for estimating the variance of the ratio of estimates because we do not have access to the covariance matrix of parameter estimates (there is one paper that does provide a covariance matrix but it is for correlations among random coefficients). Daly et al. (2012) point out that under certain conditions (generally assumed by modelers but rarely actually met), an invertible (as ours are) function of two parameters is itself a maximum likelihood estimator of the function of the true values of the parameters. We simply show that the ratio is a

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	computing WTP confidence intervals that would seem to make some misleading claims about some results reported in Daly, Hess, and de Jong (2012), who recommend using the Delta method. Carson and Czajkowski (2013) do also cite the Daly et al. paper, but only as an example of a paper that recommends using the Delta method. It is clear from their paper that they disagree with and reject this conclusion.] One reference that specifically addresses meta-analysis of WTP is by George Van Houtven (2008) ["Methods for the Meta-Analysis of Willingness-to-Pay Data: An Overview," Pharmacoeconomics 2008: 26 (11): 901-910].	first order Taylor series approximation but this extends that a bit. The Gatta et al. (2015) paper shows how this asymptotic result can produce non-trivial errors for small sample sizes. But that result is derived in the context of estimations based on surveys. In general, our survey-based estimates have large sample sizes (on the order of 1,000). It is not clear how the result applies to studies based on aggregate sales data. When one has a small survey, sampling error could be a big deal. When one's sample is, for example, the total population of car buyers, sampling error is not a concern (although the accuracy of the data may be).
	One point of contrast we can make here is that there are meta-analyses in the transport literature for another important transport-related WTP measure: Value of Travel Time Savings. I have concluded that this would seem to be a much "easier" problem than WTP for vehicle attributes, and it may be that WTP measures are measured with so much more precision that the types of problems described above do not occur. Examples are by Wardman and co-authors, e.g., Wardman et al. (2016) Transportation Research A. It would seem to be the case model estimates yield more precise estimates (in contrast to vehicle choice models), for example, outcomes that are unstable to changes in specification and/or having the incorrect sign do not seem to be much of an issue. In fact, in this literature the analyses seem to be routinely performed using ln(WTP) as the dependent variable, which does take into account skewness. At the same time, they seem to use mean values (without difficulty) rather than try to develop a more complex procedure to ascertain medians. Moreover (on a different point emphasized above),	We thank the reviewer for this additional information contrasting WTP time savings with the WTP for vehicle attributes assessment on which we are focusing. As noted by the reviewer, the problem being assessed in the value of time savings literature is in many ways less complex. The measures used are more standardized and likely better suited to the use of weighting based on sample size.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	there does seem to be attention paid to addressing variation due to differences in sample size, and there are various technical details regarding whether to treat variation due to differences in studies as fixed effects versus random effects. Moreover, there is attention paid to how results should be analyzed when multiple measures are used from the same study. (Note that I have added these remarks for the sake of completeness, to further reinforce my recommendation that the authors pay more attention to research design issues related to meta-analysis.)	
	To conclude, the material provided above in response to this charge question is primarily focused on technical issues around computing WTP-related statistics. However, it is also clear that these issues are linked to the broader context of the study itself, as addressed in the previous paragraph. This provides the basis for moving on to the next charge question.	We discuss our rationale for the statistical measures we used in the report as well as Section 1.3 of this memorandum. We believe our interpretation of the WTP values as conditional on the mean (or median) value of the price derivative is the best choice among alternatives given the available data.
	This question focuses on presentation of results, and conclusions. To review a few points from my introductory comments and the response to previous questions, the structure and content of this report can be summarized now as follows: 1. Studies were identified and selected. WTP measures were computed, largely along the lines as described in Section 4 and Appendix C. 2. Results on WTP measures and confidence intervals were summarized from individual studies in Appendix B. Summary measures across studies were presented in various forms (graphical and tabular) in Section 5.	We considered excluding non-significant estimates of attribute coefficients. In our view, the question is how to interpret non-significant estimates? In general, the attributes with non-significant estimates are relevant vehicle attributes. Frequently, other models find statistically significant estimates for the same attribute. Which finding is correct? We decided that omitting non-significant coefficients implied a different model formulation than the one actually presented in the paper. That is, had the author estimated a new model omitting the non-significant coefficients, the values of other coefficients would change in ways we could not calculate. We decided that the better approach was to take the whole model as presented in the paper but use the standard errors of the attribute derivative to characterize the uncertainty in the resulting WTP estimate. Dr. Bunch appears concerned that our choice of sample and calculations of WTP estimates may lead to an erroneous conclusion

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	3. Basic conclusions about the results include the major observation that these WTP distributions have wide variation, and frequently have the theoretically incorrect sign. At the same time, the within-study estimates of variation seem to frequently suggest more precision. 4. The apparent "lack of consensus" in these WTP results is discussed and addressed in Section 6. As part of this discussion, the authors begin to explore in more detail the possible reasons for why these values might vary across studies, and start to go into more detail by looking at more specific subsets of results from different studies.	about there being a "lack of consensus" in the WTP estimates: if we had used judgment to select only what we considered to be the best studies and most appropriate parameter estimates, and calculated the estimates using alternative methods, would the lack of consensus disappear? Perhaps it would be reduced but at the cost of introducing a very real potential for confirmation bias. And, as we have pointed out above, the data provided by almost all studies are insufficient to permit better methods for estimating the mean WTP. Our goal has been to describe what the literature has to say about WTP for vehicles attributes. As we stated, we hope that our report will encourage researchers to provide accurately calculated WTP estimates from their research, and to consider our other recommendations for advancing the state of the art. We believe that pointing out the lack of consensus will provide incentive to researchers to make these calculations.
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	At this stage, I have a couple of inter-related concerns about the status of this work. The first concern would probably be obvious from comments I provided to charge question 3. It is unclear to what degree some of the issues with the results are a function of the specific details of how the WTP distributions were computed and summarized. At the same time, the technical background I provided also sheds some light on why these measures may have potentially unavoidable difficulties due to the challenges associated with model estimation.	Recognizing that, as in any empirical analysis, there are alternative ways in which the statistical analyses could have been conducted, we believe our current methods are the best choice available given the available data. As noted by the reviewer, there are unavoidable difficulties in implementing such an analysis using the data available from peer-reviewed studies. Although the reviewer raises a number of valid considerations and topics for future research, it is unclear that any of the suggested alternatives to consider would eliminate the issues raised.
	Having said this, the technical discussion also suggests that there may be better ways to quantitatively evaluate the relative merits of results coming from different studies at an earlier stage of the process, by adhering to various procedures researchers have developed for performing meta-analysis. Specifically, in meta-analysis it would appear that measures from different studies are	As noted by the reviewer, it is possible to include various metrics to control for study quality. We limited our sample to studies published in the peer-reviewed literature and high-quality gray literature, but it is possible to weight or otherwise adjust the meta-analysis based on quality metrics or sample size. Selecting studies based on the analysts' judgment concerning the quality of the research could potentially improve the quality of the WTP estimates and perhaps reduce their variability. On the other hand, it will also introduce the analysts' confirmation biases. Because the primary goal of this report is to

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	routinely given WEIGHTS that capture at least some of these concerns.	describe the evidence on WTP for vehicle attributes in the existing literature we chose to err on the side of avoiding confirmation bias while ensuring that the studies included met the standard of professional journal peer review. As noted elsewhere, sample size varies widely in this literature and is not directly comparable in many cases.
	Now, let's digress for a moment regarding some of the current results. First, there is the problem that sometimes WTP measures have "the wrong sign." There may be no way to finesse this when faced with the authors' challenge of using existing results. But, consider the fact that estimates with the wrong sign might also frequently be measured with low precision (i.e., large standard errors). Price coefficients in vehicle choice models are notoriously difficult to get precise estimates for, and as show in some of the results, when researchers attempt to estimate price coefficients for income-based segments, incorrect signs can emerge (typically for higher income households). It may be that some judgment is required when evaluating this situation, keeping in mind that researchers were addressing multiple issues and were not solely focused on producing WTP estimates for policy analysis purposes. A similar situation would apply if a price coefficient is relatively small (but with the correct sign), leading to a very high WTP estimate. Even if the estimate is statistically significant, unless it is measured with an adequate level of precision, this will yield all of the statistical problems described in the previous section.	As noted by the reviewer, we are working with existing results from peer-reviewed and other high-quality studies. Our primary goal was to summarize what the literature has to say about WTP for vehicle attributes.
	In other words, it seems to me that a more quantitative way of discriminating across studies (e.g., using weights) should be developed, and they	As noted elsewhere, we chose not to explicitly weight the studies given inconsistency in meaning/interpretation of sample size and our

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	should be used much earlier in the process when producing results of the type that are currently presented in Section 5. (I realize that at this stage the message might be repetitive: the meta-analysis paradigm and other examples in the literature might provide a useful guide on how to do this.)	focus on summarizing the existing literature without introducing confirmation bias.
	At this stage, I need to try to "draw a line" and move on to concluding remarks. One issue here is that I want to be constructive as a reviewer, and I have tried to keep reminding myself that some of the thoughts I am generating here might be viewed as a function of "20-20 hindsight."	There are alternative ways in which the statistical analyses could have been conducted, but we still believe our current methods are reasonable and the best choice available given the available data. We encourage other researchers to explore alternative specifications for comparison with our results.
	To summarize in a more focused way: The technical issues raised in response to charge question 3 should probably be considered carefully by the authors. It may be that it is possible to use some of the theory in Carson and Czajkowski (2013) to come up with an alternative way of computing a measure of WTP central tendency and precision based on quantiles (where central tendency is the median) for the individual observations, and, if possible, that these also should be used for the purpose of developing weights. Having suggested this, it will almost certainly be the case that some type of approximation or working assumption will be required similar to the one currently made for implementing the Delta method. However, this could end up producing better results, and potentially a contribution to the literature. Next, finding a way to address the relative precision of measures from various studies and using appropriately weighted results rather than treating all results as equal would be a potential	We do not find in the C&C (2013) paper any methods that can be applied given the information generally available in the published literature. For example, the authors recommend constraining the error distributions of coefficient estimates to be lognormal or other distributions that have no support at zero. This, and the authors' other suggestions, are not possible for us to accomplish ex post without redoing all the studies. Thus, for a study such as ours, the proposed solutions are not feasible. As we explain in Section 1.3 above, the problem of a finite probability that the price derivative will be zero should be considered an artifact of the assumptions of estimation methods.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	improvement. It might be possible to make these comparisons using summary statistics.	
	This is where I recognize that I am getting even closer to the "slippery slope" of suggesting that, once the above things have been addressed, it should be straightforward to do a meta-analysis regression that would yield quantitative results to address some of the issues currently discussed in Section 6.	Meta-analysis was not within the scope of the current report, although the authors did complete another phase of the project that included meta-analysis as summarized in Greene et al. (2018).
	One final thought worth considering: we have already mentioned the difficulty of getting precision on price coefficient estimates. However, one of the primary issues of concern right now is the tradeoff in consumer preferences involving fuel efficiency versus performance. Would it make sense to do some type of exploration of how consumers directly trade off these attributes versus each other? In other words, rather than use intermediate estimates of WTP, directly compute marginal rates of substitution for these two attributes. It seems likely that the coefficients might have better precision. On the other hand: our experience is that coefficients for performance measures tend to be more unstable than for other attributes. One additional idea here is that this approach would necessarily involve a specific set of studies to compute these tradeoffs, and the differences could be compared across studies. The current situation presents results for these attributes separately, and drawing conclusions that are based on comparing the behavior of WTP estimates for each of the two attributes could be compromised by the fact that different studies are used for each separate analysis.	This is an interesting recommendation for consideration in future research.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	The current "conclusions" section includes a variety of pieces of advice for researchers going forward. Some of this is okay to say, but the authors know it would also be difficult. Researchers have lamented the lack of, e.g., "validation studies," for a very long time but there have always been practical obstacles. (However, as data availability increases this might be more reasonable.)	We agree that some of the recommendations may be challenging to implement, but the point of making those suggestions is to encourage researchers to take additional steps that we believe will improve the value and applicability of their findings.
	The authors suggest that researchers should also routinely produce WTP estimates from their models and report them to avoid the issues faced by the authors in compiling this report. Fair enough.	Agreed.
	BUT: If you are going to go down this route, you might want to include additional details that would be helpful in this regard, given the serious technical issues associated with using ratios of estimates as is done here. Specifically, note that Carson and Czajkowski (2013) directly address this type of issue, giving a suggestion for how coefficients should be estimated to avoid some of the issues associated with using ratios of coefficient estimates. However, note that their suggestion is tantamount to enforcing a strictly negative price coefficient (in a way that there is no statistical mass at zero). This gets into the potentially dicey territory of how much a researcher should impose theory-based restrictions on estimates. (This is rather subtle, since in this case the approach is more targeted toward the statistical variation in the estimate than the estimate itself.)	We agree that it is useful to add some references to the literature to our recommendations and we have made changes to that effect.

Comments by Dr. David Bunch		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	Second, I have avoided until now pointing out that the WTP discussion in Section 4 leaves out the actual underlying theory based on the conditional indirect utility function. It could be important to include some of this information in a report like this one, which can be traced back to McFadden (1981). More specific treatments relevant to this paper are, e.g., McFadden (1998) ["Measuring Willingness-to-Pay for Transportation Improvements," in Theoretical Foundations of Travel Choice Modeling, Garling, Laitila, Westin eds., Pergamon], although there are other similar papers that could be looked at.	Thank you for this recommendation. We now include the basics of the underlying theory of WTP in Section 2 of our report.
	In particular, some of the well-known results for computing closely-related measures of consumer welfare using log-sums from MNL and NL models could be worth considering. As long as researchers are being asked to compute WTP measures, why not ask them to use log-sum-based measures if they happen to be using MNL or NL? One question I have not had the band-width to explore is under what conditions these measures end up being equivalent to the coefficient ratio measure used in this paper. In any case, it is certainly likely that these measures would not have the same statistical problems as the ratio of coefficients. These are, of course, the types of measures that simply cannot be easily computed except by the researchers at the time they perform a study, because they require full access to all of the original data.	Good point. Our report is focused on marginal WTP. Logsum measures are more appropriate for estimating larger changes in relevant variables. This is worth pointing out and we have added this to Section 2.

Comments by Dr. David Bunch	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):	
At this stage, I think the approach I have taken has allowed me to cover a variety of issues in conjunction with the charge questions themselves. I continue to encourage the authors to consider finding a way to incorporate more of the principles of the "meta-analysis paradigm" in conjunction with addressing the more specific items identified above. One major implication is that the issue of evaluating the relative merits of the sources of individual WTP measures should be more formally addressed at a much earlier stage, and based on some quantitative criterion that can also be used as "weights." [Bunch has also posted the following comments outside of the original table. ICF has copied in here.] The following material constitutes my review and comments. The primary structure followed is as prescribed in the instructions we were given in the charge letter, i.e., to generate responses to specific Charge Questions in tabular form. To improve the efficacy of the process of generating these responses as well as readability, I have taken the liberty of slightly revising the tabular format. In addition, I would add the following: There are aspects of the review request discussed in the charge letter that appear in the paragraphs preceding the charge questions themselves. I consider them to be 1) important and relevant, but 2) not necessarily adequately reflected in the Charge Questions themselves. Although it might have been possible to postpone this discussion to the last sections of the table, I decided that I wanted to instead include some introductory material prior to answering the specific questions. Some of the issues were also explored in more depth during our kickoff conference call, where we posed a number of questions and discussed some preliminary reactions to the report. First, here are some comments in response to specific excerpts from the charge letter. "EPA has been developing estimates of the willingness to pay for (WTP) a variety of vehicle characteristics, such as fuel economy, performance, and comfort. If vehicle standards that EPA issues affects these other vehicle characteristics, having estimated values for these attributes might enable EPA to develop better estimates of the benefits and costs of its standards. … The estimates are based on existing published studies of light-duty vehicle demand in which authors have	We have taken on board the recommendation to include a description of the principles of meta-analysis that we followed. However, we remain convinced that our criteria for selecting WTP estimates are appropriate given the purpose of our study.

econometrically estimated purchase patterns for vehicles based, in part, on vehicle characteristics." Before responding to the charge questions more directly, it will be helpful to review at a very high level what this report consists of. The work basically consists of: 1. A systematic literature review consisting of identifying and collecting studies that perform data analyses and/or model estimation yielding results that can be used to compute WTP estimates for various vehicle attributes. 2. An explanation of methods used for computing WTP estimates. 3. Summarization of computed WTP statistics and distributions for various individual attributes, at both the level of individual studies, and compiled across studies. 4. Discussion related to what was observed in the results. This was primarily focused on the fact that WTP estimates appear to exhibit large variability across studies. Some of this discussion involved considering in more detail specific results from selected studies to explore possible reasons for this variability, attempting to evaluate the possible impact of various research design dimensions (e.g., different data types, model types, estimation approaches, etc.).	
Again, the key overall finding (that WTP estimates from the vehicle choice modeling literature appear to vary widely, and even take on theoretically incorrect signs) is unsurprising and has been an ongoing concern of researchers and policy makers alike for quite some time. We discussed this problem/issue during the kickoff call, because it would seem to beg the question of whether or not this type of project should not have been more concerned with trying to evaluate the relative quality of the studies, rather than simply collecting and reporting numbers. To clarify, the charge letter itself does include the following: "We ask that you review methods and underlying assumptions, their consistency with the current science as you understand it, and the clarity and completeness of the presentation. For this review, no independent data analysis is required. Rather, we ask that you assess whether the data and methods are applied appropriately, given the state of current understanding, and the conclusions reasonably drawn."	We agree that our findings are not likely to be a surprise to practitioners in this area. However, to our knowledge there does not exist a documentation of that understanding approaching the scope and thoroughness of this report.
We asked for, and received clarification that this part of the request primarily applies to the "methods and underlying assumptions" of the review itself, not the underlying studies on which the review is based. However, what appears next is:	This is an important point. We did not interpret our task to be a critical evaluation of the modeling methods and data in the literature. Instead, we interpreted our task to be discovering what the existing literature implied about WTP for vehicle attributes. If the literature produced a robust

"Note that EPA's interest in these estimates is based not only in seeking robust estimates for values of willingness to pay for vehicle characteristics, but also in understanding the robustness of the models from which these estimates are derived."	consensus, that would be our finding. However, as a general rule, it did not, and that was our finding.
Note that this expands the scope of EPA's interest to understanding "robustness of the models from which these estimates are derived," in order to add clarity to the relative robustness of the estimates themselves. This was a subject we attempted to explore during the phone call. During this discussion, we suggested/asked whether or not this study shouldn't go more in the direction of "meta-analysis," rather than stop short by limiting the review to the collection of data and reporting of summary statistics. In either case, it seems that finding a way to evaluate the relative merits of various "sources" of "WTP estimates" is potentially important to EPA.	As explained in Section 1.1 of this memorandum, we did follow the generally accepted procedures for meta-analysis during the course of the study because we were planning ahead for conducting a meta-analysis in a second phase of the study. We now explain this more clearly by referencing the steps outlined in van Houtven (2008).
These are the issues that I found myself concerned about finding a way to address, in an attempt to provide something of value to EPA in my review. To this end, I spent some time researching the subject of meta-analysis (which has not been a specific area of specialty for me). One thing I concluded through this effort is that it would be incorrect to view "doing a meta-analysis" as an extension of the steps already taken in this review.	We do not agree with this assessment given that we were preparing to conduct a meta-analysis in a subsequent phase of this project while preparing the report. The resulting Greene et al. (2018) paper is currently undergoing a third round of review at a peer-reviewed journal.
Rather, I concluded that the steps performed for this review should actually be viewed as corresponding to the initial steps within the overall process of "doing a meta-analysis." In particular, designing the procedures for identifying and collecting studies, creating rules for determining which studies to include or exclude, specifying the technical details for computing measurements and relevant statistics for "Effect Sizes" (or, "ESs," which in this case are WTPs for attributes), deciding on a rule for eliminating outliers, and even producing summaries of Effect Size distributions represent the steps of a meta-analysis that then typically culminate in building models for the ES data that seek to identify and clarify the specific sources of "variation" across studies.	We have addressed our rationale for not using sample sizes to weight estimates in Section 1.2 of this response to comments. In addition, carrying out a meta-analysis to produce central tendency WTP estimates was not the goal of this report. Rather, its goals were to produce a comprehensive database of estimates for the United States based on the current literature and to describe the nature of those estimates.
To be clear, there is usually a process that also weights the relative importance of data points from different studies using the inverse of a measure of "internal variance" for the study, based on whatever the most appropriate statistics are for the situation being studied. These variance measures are typically a function of sample size, e.g., studies with small sample size are expected to yield more biased and/or noisy estimates and therefore should be "down-weighted." Moreover, these weights are also frequently employed when producing the reported summary	As noted elsewhere, it is possible to include various metrics to control for study quality. We limited our sample to studies published in the peer-reviewed literature and high-quality gray literature, but it is possible to weight or otherwise adjust the meta-analysis based on quality metrics. Selecting studies based on the analysts' judgment concerning the quality of the research could potentially improve the quality of the WTP estimates and perhaps reduce their variability. On the other hand, it will

statistics (such as those appearing in this review). All of these steps, taken together, are important aspects of designing a meta-analysis that addresses the needs and goals of the research. For example, even the first step of defining how the universe of studies is determined should be based on these considerations.	also introduce the analysts' confirmation biases. Because the primary goal of this report is to describe the evidence on WTP for vehicle attributes in the existing literature we chose to err on the side of avoiding confirmation bias while ensuring that the studies included met the standard of professional journal peer review. We acknowledge that by including some studies from the gray literature we have introduced our own judgment about whether they meet that standard. However, including studies from the gray literature is also frequently recommended in designing meta-analyses to mitigate "publication bias" (van Houtven, 2008, p. 904). We are encouraged by the fact that none of the reviewers suggested that any of the gray literature studies we included should not have been included
I wanted to report this, since at the time of our phone call I tried to articulate a general concern about whether or not relative "study quality" should somehow be addressed in light of EPA's interests, and, if so, how that might be achieved. I believe this additional background on meta-analysis sheds some light on this question. The above remarks represent an overall sentiment that I will keep coming back to in the more detailed comments that follow.	As noted elsewhere, it is possible to include various metrics to control for study quality. We limited our sample to studies published in the peer-reviewed literature and high-quality gray literature, but it is possible to weight or otherwise adjust the meta-analysis based on quality metrics. Selecting studies based on the analysts' judgment concerning the quality of the research could potentially improve the quality of the WTP estimates and perhaps reduce their variability. On the other hand, it will also introduce the analysts' confirmation biases. Because the primary goal of this report is to describe the evidence on WTP for vehicle attributes in the existing literature we chose to err on the side of avoiding confirmation bias while ensuring that the studies included met the standard of professional journal peer review. We acknowledge that by including some studies from the gray literature we have introduced our own judgment about whether they meet that standard. However, including studies from the gray literature is also frequently recommended in designing meta-analyses to mitigate "publication bias" (van Houtven, 2008, p. 904). We are encouraged by the fact that none of the reviewers suggested that any of the gray literature studies we included should not have been included
ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:	
Given the specific direction that this review took, I have elected to forgo more detailed and specific items in individual sections.	NA

Comments by Dr. David Brownstone		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	Almost—they appear to have done an excellent job finding all of the relevant studies—at least I am not aware of any that should be included. However the key database they construct (Appendix B) fails to note the population used to construct the WTP estimates for each study. For example some studies use market data and therefore include all light vehicle sales—including rental cars, commercial fleets, etc. Other studies use survey data that may only cover new personal vehicle purchasers or possibly a representative sample of some geographic area. If there is heterogeneity in WTP across the US population, then different WTP estimates could be partially explained by different underlying populations.	Thank you. As noted in Section 1.1 above, we have taken steps to augment our description of the populations included in each study, with additional summary statistics added in Section 3 of the report. Previously we included only time period and geography. We have added whether the study includes new car buying households, new and used car buying households, all households, or all vehicle purchasers.
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	The report is clear about the methods used to calculate WTP and the "confidence bands" around these estimates, but they are not clear about how to handle models that include heterogeneity across the target population. Ideally they should compute the population average WTP and a confidence band around this quantity for each study. While this may not be possible for some studies, it would be better if this was clearly elucidated as the goal of the exercise. This would make population heterogeneity a secondary issue, but currently the report sometimes uses heterogeneity arising out of random parameters to represent estimation uncertainty.	We described population heterogeneity for all the random coefficient models based on ±1 standard deviation of the estimated distribution of the random attribute coefficient. For these models we also calculate a measure of uncertainty of the mean or median WTP estimate based on ±1 standard error of the central tendency measure of the attribute coefficient. However, we do not show both measures for the random coefficient models but focus instead on the heterogeneity measures. We have tried to make this clearer in the revised report.
3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are	The report does an excellent job (in Appendix C) of describing the issues with getting unbiased estimates of WTP from standard choice models. They also give a reasonable formulas for the bias and variance of this WTP measure in equations C-1 and C-5. Unfortunately they then ignore these	It is correct to say that we have underestimated the confidence bounds of the mean of the ratio of the attributes and price derivatives. This follows from our decision to interpret our WTP estimates as conditional on the central tendency estimate of the price coefficient

Comments by Dr. David Brownstone		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	formulas when computing their "confidence bands" as described in the middle of page 4-4. The "standard errors" they calculate are always too small. A better approach would be to calculate the bands using equations C-1 and C-5 assuming the bounds given by $-\sigma\alpha\ \sigma\beta \leq \sigma\alpha\beta \leq \sigma\alpha\ \sigma\beta$. The endpoints of this interval can be estimated and put into formulas C-1 and C-5 to generate valid bounds for the bias and variance of the WTP estimates. For example in a recent study I calculated WTP = 4 with a correct standard error of .09. The bias in this WTP estimate using equation C-1 lies between 0 and .02. The approach used in the report gives a standard error estimate of .012 (a very big underestimate), while the using the bounds on $\sigma\alpha\beta$ in equation C-5 imply that the correct standard error lies between 0 and 0.4. Unless the bias in the WTP estimates using the above bands are small, then the tables and figures should be redone showing the bands around the WTP estimates. Likewise the "confidence intervals" should be redone using the bounds.	and the use of only ±1 standard error. This is explained in greater detail in Section 1.3. The approach proposed by Dr. Brownstone is indeed a valid method for estimating the maximum uncertainty bounds based on standard errors. This can be readily seen from the definition of the correlation coefficient: $\rho(x,y) = Cov(x,y)/(\sigma x*\sigma y)$ Since $-1 \leq \rho(x,y) =< 1$, setting $\rho(x,y) = -1$ and solving the first equality for $Cov(x,y) = -(\sigma x*\sigma y)$, then doing the same for $\rho(x,y) = 1$ to get $Cov(x,y) = +(\sigma x*\sigma y)$, produces upper and lower bounds. However, these would be bounds for the mean rather than our conditional mean metric. In addition, since we don't know what $\rho(x,y)$ is, we don't know if the true bounds are close to the maximum we have calculated or far away. Our method accurately describes the error in the attribute derivative, assuming the value of the price derivative as a given. The proposed method gives an upper bound in every case that may or may not correspond well to the actual error bounds. The proposed method is useful if the goal is to understand the maximum possible uncertainty bounds. However, that is not our intention.
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	The overall conclusion is that there is too much variability in the existing estimates to find any useful "consensus" values. Since the "confidence bands" computed in the report are always too small (see 3. above) the key conclusion follows from the results.	Thank you; we agree with this assessment.
	The report offers some reasons for the lack of accurate WTP estimates, and these are generally reasonable. Since the purpose is to estimate WTP I am surprised that they did mention estimating the models in WTP space as advocated by Kenneth Train (see reference below). This greatly simplifies	Thank you for pointing this out. We address this in Section 1.3 above: only one of the studies in our study estimates its coefficients in WTP space.

Comments by Dr. David Brownstone		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	computing confidence bands for the WTP estimates and also makes it easier to model heterogeneity in WTP while still enforcing sign restrictions.	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):		
This report is based on an exhaustive review of the recent US-based literature on the willingness to pay (WTP) for different vehicle attributes. These WTP measures are typically compared to cost estimates to provide an additional unit of the attribute, but it is not clear how this corresponds to behavior in the real vehicle marketplace. Indeed many of the papers reviewed by this report were not designed to "measure" WTP, but instead to predict demand for an exogenously-specified set of real and hypothetical vehicles. Near-term policy evaluation might be done better using scenario forecasts with realistic vehicles. For example instead of looking at the difference between WTP and the engineering cost of a 1c/mile reduction in operating costs, the vehicle choice models could be used to discover the price subsidies needed to achieve a target market share for a realistic set of more efficient vehicles. Alternatively the welfare loss of achieving a particular fuel economy goal could be measured.		The reviewer raises valid points that are addressed briefly above in Section 1.4 of this response to comments. We developed a number of recommendations based on the results of this analysis.
I will take the goal of measuring WTP as given for the rest of this review. If authors of the papers reviewed were asked to estimate WTP now, I suspect that many of them would reformulate their models to work in WTP space (Kenneth Train, Riccardo Scarpa and Mara Thiene, Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps, American Journal of Agricultural Economics, Vol. 90, No. 4, pp. 994-1010, 2008.). This would allow for directly modeling population heterogeneity in WTP. More importantly it would make it much easier to compute consistent confidence intervals to represent the uncertainty in the WTP measurements.		We have included the suggestion that authors consider working in WTP space in the recommendations from our report, as described in Section 1.4 of this response to comments.
ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:		
Figure 5.1 These figures are hard to read—especially in black and white. I suggest representing each study by a vertical line extending from the top to the bottom of the "Confidence Interval" with a bold dot on the point estimate (similar to Figure 6-1 on page 6-6 in the report). These lines could be arrayed from low to high point estimates as in the current graphs.		We will make the report available in pdf format, in color.

Section 5.1.3—the claim that air conditioning is "perhaps inconsequential" since it is almost standard equipment is not correct for studies looking at battery electric vehicles. The use of air conditioning seriously degrades vehicle range for these vehicles, leading to a potential demand for new technologies like better vehicle thermal insulation and (possible solar-powered) vehicle ventilators.	Thank you for pointing this out. We have removed the incorrect statement from the report.
Page 5-14—the claim on the top about the differences between stated and revealed WTP estimates does not seem to be supported by the values in Table 5.2. See also second to the last paragraph on page 6-1.	Thank you for pointing this out. We have corrected the statement. It is the median estimates based on stated preference and revealed preference data that differ.

Comments by Dr. George Parsons		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	The Literature Review (Section 2) and Description of Studies and Attributes (Section 3) are clear, complete, and systematic. The coverage of studies is good and the method used to identify and include studies is good. The box on page 3-1 and supporting text are a good description of the selection process. The studies include the best current and recent-past research as far as I could tell. The direct contact with authors, which apparently uncovered additional studies, is impressive and makes this more thorough than most investigations of this sort. The tabulation by year and study type is also very helpful in understanding the data. I have a few suggestions:	Thank you; detailed comments are addressed below.
	(1) It would be useful to see an overlay of the 52 "main sample" studies on figure 3-1 to see where the final selections fall over time. (2) It would be useful to see a frequency distribution for the top 5 (maybe 10) journals represented in the sample and their percent frequency in the sample. (3) In addition to a paper count, it would be interesting to see a unique data set count, since some of the 52 studies use the same data. Sort of a Table 3-1 for data sets.	Thank you for these suggestions. We have incorporated additional summary data into Section 3 of the report to help address some of these issues.
	(4) Somewhere early in the study, we should be told more about how EPA will use the numbers—maybe an exploratory example application. I kept wondering how will these many different points estimates would be used to get at welfare effects. I know EPA does an effort like this for VSL's (i.e., summarizing literature to get at a VSL to use in policy analysis). In that case, we know you will be	To the best of our knowledge, EPA has not made a final decision about how it will use the results of this study except to generally inform their decision making.

Comments by Dr. George Parsons		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	multiplying the VSL by number of lives saved or lost. It is easier. But here, the welfare effects come through market choices in response to a policy (fuel standards). It is more difficult. Having some idea upfront of how the numbers will be used, would help. I might be missing something obvious here! Overall, the presented description is sufficient to form a general view of the quality and quantity of the data used.	
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	The Methodology (Section 4) is mostly clear and complete. The description of how the "central tendencies of WTP" are calculated is straightforward and the discussion on why the first order approximations are used is good. I have several suggestions here (I number these beginning where I left off above for easy reference):	Thank you; detailed comments are addressed below.
	(5) On page 4-1 (point 3 near the top), there is reference to "…other models with random distributions…" What other models are these? They are never described or discussed as far as I could tell.	These are BLP models (using the method of Berry, Levinsohn and Pakes, 1995) and variations thereon (e.g., Petrin, 2002). It is chiefly the estimation method and use of aggregate sales data that makes these models different from Mixed Logit models. We have added text to make this clearer.
	(6) On page 4-3 (second full paragraph), for NMNL models you write "…the utility functions of the nests that include the prices of vehicles are used in estimating WTP…" Is it possible to have more than one value for an attribute from a given study because price-attribute pairs show up in more than one nest with different coefficient estimates or different scale estimates? If so, how was this handled?	Yes, this is possible and it happens within our main sample. In that case we calculate WTP measures for each nest.

Comments by Dr. George Parsons		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	(7) You do not mention cases where price might enter non-linearly (e.g., ln(Pr)). Are there any of these and, if so, how are they handled?	Yes, there are such cases. We always divide the derivative of the utility function with respect to an attribute by the derivative with respect to price. In the case of bln(P) the derivative would be b/P. In that case we would use the mean price from the sample in question. We have added text to clarify our handling of such cases.
	(8) On page 4-3, you write that they use "…mean values for normally distributed random coefficients and median values for lognormally distributed coefficients…". I assume you use means for triangular, uniform and other distributions that show up in the literature, but nothing was stated, so we cannot be sure. Aside from these points, the description of the methods is sufficient to form a general view of the quality and quantity of the WTP calculations.	That is correct, we use medians only in the case of lognormally distributed coefficients. We have attempted to clarify this point in the text.
3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	I found no technical errors but do have a couple questions about reasonableness of the procedures.	Thank you; detailed comments are addressed below.
	(9) As I read the manuscript, you use simulation to estimate a dispersion of WTP estimates from mixed logit models but do not use it for computing means. Why not simulate to calculate unbiased means?	We do not use results from simulations of MXL models unless the author provides them and then we use the results of authors' simulations. We did not do any simulations of our own. We have attempted to clarify our methods in the text to avoid this misunderstanding.
	(10) Given that the automobile market is highly competitive, I believe the models based on RP data largely sketch out marginal cost, not demand functions. Think about variation in the price of a car due to shoulder room. Adding x square feet of shoulder space comes at an added cost of Δc to the manufacturer. Competitive forces, in theory, should force the price difference for a small should-room car versus a large shoulder-room car (all else constant) to be near Δc, so $\Delta p = \Delta c$—a	This comment applies to hedonic price models but not to random utility models of vehicle choice. There are seven hedonic models in our sample. They do not all make use of methods to identify the demand function, so your comment may well apply to them. However, we have made no attempt to correct these studies and we include them since they meet our criteria. We agree that the central tendency estimates we derive may be local values that may not apply for non-marginal changes. We have added text noting that changes over intervals can be calculated using logsums.

Comments by Dr. George Parsons		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	standard result from micro-theory. This should hold for all attributes over all levels. The price differences, in principle, then capture marginal cost differences in RP data. (The classic identification problem in the hedonic setting is that RP data sketch out an implicit price schedule and not demand or supply—true, but in a case of competitive market like autos, that price schedule will be a cost function.) As far as WTP goes, you will have an estimate of marginal WTP at the equilibrium implicit prices for different consumer groups, but not the demand function. I think this is important to keep in mind when analyzing the results. Since the primary interest is WTP and ultimately non-marginal changes in WTP, since changes in fuel standards and any ensuing changes in comfort, performance, etc., are likely to be non-marginal, you need to be careful. For this reason, I give the nod to SP data in this setting—despite its being hypothetical. The price changes paired with the varying attribute levels in the SP case are not market-based. Instead, they are set by researchers, and should in principle be sketching out demand functions and hence WTP for non-marginal changes. This comment falls in the "reasonable disagreement" category. Otherwise, I thought the procedures were appropriate and reasonable.	
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	The results are presented mostly in an appropriate and reasonable way. Section 5 was logical and clear. I noticed the discussion of Vehicle Class was missing. Also, the section could perhaps benefit from a more consistent use of graphs and tables across attributes, but that is not major. I	We agree that the vehicle class WTP estimates are generally not comparable and have removed them from the table for the reasons you describe below. Not only do different studies use a different default class but the numbers and definitions of classes vary, as well. Thus, it would require a substantial amount of subjective judgment to provide consistent estimates. We have added text to make this clearer.

Comments by Dr. George Parsons		
CHARGE QUESTION	**COMMENTS**	**AUTHOR RESPONSES**
	suppose some deserve more attention than others. I do have a concern about how Fuel Type and Vehicle Class are interpreted.	
	(11) The WTPs for Fuel Type and Vehicle Class come from Group Specific Constant coefficient estimates, which must be relative to some excluded group. For Fuel Type I am assuming the excluded group is a conventional gasoline vehicle. It is never clearly stated or discussed in the text, but gas is excluded in Table 5-1. Since many of the key attributes that distinguish gas from electric and other vehicles are included in the models, the Group Specific Constants are picking up average "unobserved" differences between gas and electric (or whatever other fuel type is in question). Of course, most of what is driving the difference in gas versus electric is likely to be included in the variables like range, fuel time, performance, etc., so the Group Specific Constant WTPs are picking up the residual differences only. Indeed, if you included all attributes that matter to people, you would expect these constants to be near zero. In any case, the text does not seem appreciative of meaning of these coefficients and their WTPs. I may be misreading the text but they read more like they are interpreting the WTP as picking up the full difference between a gas and electric vehicles and not a residual unobserved effect. Some refinement is needed here. Vehicle Class has the same issue. In this case, I am also wondering how they kept this exclude group (whatever it is) constant across the studies.	This is a good point. In many models, vehicle class WTP values depend on what the excluded or reference class was. Although we calculate these WTP estimates, they are generally not comparable across models because vehicle class definitions and reference classes both vary from study to study. As a result, we do not present summary statistics for WTP for vehicle classes. The reviewer is correct that the excluded group for fuel type is always a conventional gasoline vehicle. We have added language to that effect to the report.

Comments by Dr. George Parsons	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):	
I have a few suggestions for simplifying the analysis and directing it more toward its final use in policy. As it stands now, it is just "too busy" and needs some researcher judgement calls to make it more manageable and useful for policy. I know the wide range of values makes application seem almost futile but EPA needs to proceed somehow and I think the literature has value. Here are my suggestions:	Thank you for providing these useful recommendations, which we will consider for future research. We believe our analysis makes a useful beginning from which to consider how to improve WTP estimates and potentially develop a greater consensus.
(12) Focus on the key set of attributes that EPA will use for policy analysis. Clearly fuel savings and some of the performance measures stay, but narrowing the set would make it easier to digest and bring focus on to what matters. Do we really need to look at air conditioning, fuel availability, auto-transmission, and so on?	In our meta-analysis, now mentioned in this report, we focus primarily on fuel cost and performance. We nevertheless have included these other variables, because they may contribute to our understanding of the role of specification (e.g., omitted variable bias) in the results.
(13) Instead of using means for individual characteristics from the sample where each study was done to compute WTP, consider creating some representative consumer types and estimate values for each of these from each chosen study. The groups might be defined by income, family size, and location.	In our view, (13) is another valuable recommendation for future research. Namely, how can the results from studies that provide different WTP estimates for population subgroups be used to describe the heterogeneity of preferences in the car-buying population?
(14) Consider using one WTP outcome per paper, a sort of preferred model. Having many estimates from one paper or one paper dominating the set of studies for one attribute is misleading. Just because an author of a study considers another model, they get greater representation. You could even consider one model per data set.	Comment 14 is discussed in Section 1.5 above. The reported non-preferred outcomes provide insights into the robustness of the primary results and may contribute insights into sources of variation. In addition, not all papers identify which model is preferred.
(15) Focus on SP studies for reasons I mentioned above.	This comment (15) is addressed towards the end of Section 1.3 above.
Finally, in a somewhat different but related vein, I also think it is worthwhile to consider some "directed" simulations using just a handful of studies: (16) These would be selected based on the quality of theory, method, and data used and capacity to perform realistic analysis (i.e., include a useful array of attributes). Then, consider estimating welfare effects for non-marginal changes in several attributes simultaneously using these model. These would be realistic, policy-relevant scenarios meaning to capture EPA concerns (e.g., fuel economy improves, performance drops, and comfort drops). These might use a log-sum calculation, in which case the analyst would have to somehow formulate a representative choice set. Honestly, I have not thought through the details and may	(16) We agree that attempting to select studies and estimates by judging the quality of the methods and data is a valid approach. However, we have described the reasons we took a less subjective approach in Sections 1.1, 1.2, and 1.5, above.

be missing potential pitfalls, but I like the idea of treating some studies as better than others and choosing them for application and getting values of direct use for policy. This approach also allows for substitution, which is likely to be critical for getting at the true welfare effects—something separate point estimates of price do not. Another strategy is to do your own primary analysis using SP data and targeting the specific policy questions at hand. You could even do some contingent behavior analyses to get at specific issues of policy relevance. This is probably you first best solution.	
ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:	
NA	NA

References

1. Berry, S., J. Levinsohn and A. Pakes, 1995. "Automobile Prices in Market Equilibrium," Econometrica, vol. 63, no. 4, pp. 841-890.
2. Busse, M.R., C.R. Knittel, and F. Zettelmeyer. "Are Consumers Myopic? Evidence from New and Used Car Purchases." American Economic Review 103, no. 1 (2013): 220-256. doi:10.1257/aer.103.1.220.
3. Carson, R.T., Czajkowski, M. 2013. "A New Baseline Model for Estimating Willingness to Pay from Discrete Choice Models", International Choice Modelling Conference, Sydney, Australia.
4. Daly, A., S. Hess and G. de Jong, 2012. "Calculating errors for measures derived from choice modelling estimates", Transportation Research Part B, vol. 46, pp. 333-341.
5. Gatta, V., E. Marcucci and L. Scaccia, 2015. "On finite sample performance of confidence intervals methods for willingness to pay", Transportation Research Part A, vol. 82, pp. 169-192.
6. Greene, D.L., A. Hossain, J. Hofmann, G. Helfand, R. Beach, 2018. "Consumer Willingness to Pay for Vehicle Attributes: What Do We Know?", manuscript under peer review, RTI International, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709-2194.
7. ICF, 2017. Peer Review of "Consumer Willingness to Pay for Vehicle Attributes" –Final Report, EPA Contract No. EP-C-16-020, Work Assignment No. 1-16, San Francisco, CA, November 15.
8. Loomis, J.B., 2014. "Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys", Journal of Agricultural and Resource Economics, vo. 39, no. 1, pp. 34-46.
9. Train, K. and M. Weeks. (2005). "Discrete Choice Models in Preference Space and Willingness-to-Pay Space." In: Scarpa, R. and A. Alberini (eds), *Applications of Simulation Methods in Environmental and Resource Economics, The Economics of Non-Market Goods and Resources*, vol. 6: 1-16. Springer: New York. doi:10.1007/1-4020-3684-1_1.
10. van Houtven, G., 2008. "Methods for the Meta-Analysis of Willingness-to-Pay Data", Pharmocoeconomics, vol. 26, no. 11, pp. 901-910.

Peer Review of "Consumer Willingness to Pay for Vehicle Attributes" – Final Report

EPA Contract No. EP-C-16-020

Work Assignment No. 1-16

November 15, 2017

Submitted to:
U.S. Environmental Protection Agency
2000 Traverwood Drive
Ann Arbor, MI 48130

Submitted by:
ICF
620 Folsom Street, Suite 200
San Francisco, CA 94107

Table of Contents

I. Introduction 3
II. Peer Review Process 3
1. Selecting Reviewers 4
2. Administering the Review and Receiving Comments 5
3. Difficulties Encountered 5
III. Responses to Charge Questions 5
1. Comment Overview and Summary 5
1.1 Is the Choice of Publications Described Sufficiently 6
1.2 Are the Methods Described Sufficiently 7
1.3 Are the Methods and Procedures Technically Appropriate 7
1.4 Are the Results Presented Appropriately 8
1.5 Additional Comments 10
2. Comments by Reviewer 12
2.1 Comments by Dr. David Bunch 12
2.2 Comments by Dr. David Brownstone 31
2.3 Comments by Dr. George Parsons 35
Appendix A. Resumes of Selected Reviewers 42
Appendix B. Conflicts of Interest Declarations 71
Appendix C. Peer Reviewer Mid-review Meeting Notes 75

I. Introduction

EPA has been developing estimates of the willingness to pay (WTP) for a variety of vehicle attributes, such as fuel economy, performance, and comfort. If vehicle standards that EPA issues affect these or other vehicle attributes, having estimated values of the WTP of customers for these attributes would enable EPA to develop better estimates of the benefits and costs of its standards. "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge" (the Report) documents the development of these WTP estimates for a wide variety of vehicle attributes. The estimates are based on existing published studies of light-duty vehicle demand in which authors have econometrically estimated purchase patterns for vehicles based, in part, on vehicle attributes. In these studies, it is often possible to derive estimates of WTP from the econometrically estimated demand models, even in cases when the authors of the studies do not publish these estimates in their papers.

EPA's Peer Review Policy Statement, 2006, states that external peer review is expected for highly influential scientific assessments, and is the "approach of choice" for influential scientific information or for work products that have special importance. Therefore, EPA submitted the Report for peer review, seeking the reviewers' expert opinion on the methodologies employed and analyses presented in the report. ICF facilitated this peer review, and this memorandum contains a summary of the peer review results as well as documentation of the peer review process.

This project was conducted from August to November 2017. The peer review process followed the current version of EPA's *Peer Review Handbook.*[1] This document presents the unedited comments and conclusions presented by each peer reviewer along with a brief summary by charge question. The curriculum vitae (CV) and conflict of interest (COI) statements for each peer reviewer are provided in Appendices A and B.

The following information is provided in this Technical Report.

1. Description of the peer review process (Section II)
2. Reviewer responses to charge questions (Section III)
3. Reviewer supporting documentation (Appendix A and 0)
4. Notes from mid-review meeting with EPA, ICF, and the contracted peer reviewers (Appendix C).

II. Peer Review Process

Consistent with EPA's *Peer Review Handbook,* ICF conducted the peer review in three stages. We first identified a qualified set of reviewers. ICF then contracted with all reviewers, conducted the review, and collected reviewer feedback on the report. Finally, ICF documented the peer review process, consisting of this report, to submit the assembled information from the peer

[1] U.S. Environmental Protection Agency, Peer Review Handbook, 4th Edition, October 2015. Prepared for the U.S. EPA by Members of the Peer Review Advisory Group, for EPA's Science Policy Council, EPA/100/B-15/001. Available at http://www.epa.gov/osa/peer-review-handbook-4th-edition-2015-0, including OMB's Information Quality Bulletin for Peer Review (Handbook, Appendix B) provisions for the conduct of peer reviews across federal agencies.

reviewers to EPA. Ultimately, EPA will convey results of the peer review process to authors of the Report, who will respond to the comments.

The following sections provide detail on these steps.

1. Selecting Reviewers

ICF first identified a pool of independent subject matter experts from which to select three peer reviewers that represent the best qualified candidates. Qualifications included two technical considerations. The first is any actual or apparent conflict(s) of interest or lack of impartiality that would preclude an independent review. The second is that the combined expertise cover the two focus fields of this analysis:

1. Academic literature on vehicle demand modeling; and
2. Derivation of estimates of WTP from consumer demand models, including discrete choice and hedonic models.

ICF identified twelve potential reviewers for the report based on a combination of individuals originally suggested by EPA and those identified through our own research. ICF then contacted each candidate by e-mail and/or telephone and ascertained their availability and qualifications to perform the peer review within the allotted schedule. This contact was designed to assess each potential reviewer's expertise in the field, their ability to perform the work during the period of performance, any association they have with the work that would preclude them from being independent and reasonably expected to be objective. We also collected a CV or resume from each peer reviewer who expressed an interest in participating.

Based on these contacts, ICF selected three qualified independent reviewers to conduct the peer review, with the goal that the combined expertise of the selected reviewers would cover all technical aspects of the report. ICF suggested the following reviewers in our August 14, 2017 Final Peer Review Selection Memo to EPA:[2]

1. Dr. David Bunch
   Professor of Management, Graduate School of Management
   University of California at Davis
   One Shields Avenue
   Davis, CA 95616
   dsbunch@ucdavis.edu
   (530) 752-2248

2. Dr. David Brownstone
   Professor of Economics, Department of Economics
   University of California at Irvine
   3151 Social Science Plaza
   Irvine, CA 92697
   dbrownst@uci.edu
   (949) 824-6231

[2] Draft peer reviewer selection memo (Task 1) Contract EP-C-16-020, Work Assignment No. 0-16: Peer Review of "Consumer Willingness to Pay for Vehicle Attributes", to Gloria Helfand, US EPA OTAQ, from: Jon Hecht and Andie Fritz, 14 August 2017.

3. Dr. George Parsons
   Professor, School of Marine Science and Policy
   University of Delaware
   204 Robinson Hall
   Newark, DE 19716
   gparsons@udel.edu
   (302) 831-6891

These selected reviewers each possess the experience and technical expertise required to conduct the review. ICF anticipated that all relevant areas of the peer review were sufficiently covered based on this selected group of reviewers. ICF's Peer Review Selection Memo documented this process. EPA concurred with all selected reviewers.

2. Administering the Review and Receiving Comments

ICF composed and delivered a charge letter to the three selected reviewers, along with the "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge" report. The charge letter included instructions on how to complete the review, a timeline of when comments were due to ICF, and a conflict of interest (COI) form. ICF sent the charge letters to Bunch and Brownstone on September 7, 2017, and to Parsons on September 13, 2017.

ICF then arranged and hosted a mid-review teleconference on September 21, 2017 between the selected peer reviewers, EPA, and ICF staff. During this 1-hour meeting, the peer reviewers and EPA staff responsible for the analysis were introduced. The meeting included extended discussions on the nature of the review, background information on the review itself, and technical issues for consideration. ICF's notes from these meetings are attached to this report as Appendix C.

ICF requested the peer reviewers provide comments within two weeks. All peer reviewer comments, cover letters, and completed COI forms were received by ICF by October 6, 2017.

3. Difficulties Encountered

No significant difficulties were encountered while performing this review.

III. Responses to Charge Questions

Section III.1 presents a brief summary overview of the feedback received on the four charge questions. This is followed by Section III.2, which provides the direct, unedited peer reviewer responses to each of the charge questions. Within this section, the detailed responses are organized by reviewer, with responses to each of the applicable charge questions. The responses are all presented in the same table format, with the leftmost column listing the charge question, and the middle column providing the reviewer comments.

1. Comment Overview and Summary

The following overview of the peer reviewers' comments to the specific charge questions is a brief summary of the broad, general themes among the comments for each question. It does not

rewrite the responses or supersede the more detailed, direct comments presented in Section III.2.

All three reviewers provided other and additional comments beyond those of the four prescribed questions, which are summarized in Section III.1.5.

The questions have been abbreviated for presentation here.

1.1 Is the Choice of Publications Described Sufficiently

Generally, the reviewers thought that the choice of publications was described sufficiently, with both Brownstone and Parsons noting that the choice of studies appeared to be thorough. Parsons also noted that the descriptions of the studies and attributes included in the analysis were clear and sufficiently descriptive.

Bunch had some suggestions for this area. He noted that there was not enough detail on how studies were selected for inclusion. He felt it would have been helpful to describe the selection criteria and process because some studies that he was aware of were excluded, and he was not sure of the reasons for their exclusion. He commented that the selection process could have been structured more like the selection process for a meta-analysis, in terms of rigor.

Bunch also noted that he was not certain that some of the selection criteria was appropriate, given the objectives of the study. In particular, he discussed the decisions to only include US-based studies and studies conducted after 1995. He commented that on one level these restrictions are logical, but on another level they could result in the exclusion of studies that would be useful and that would add to the statistical power of the analyses, which often suffer from small sample sizes. In summary, Bunch noted that the current version of the study makes it difficult to assess the quality and quantity of the data used from the included studies.

The reviewers offered suggestions to improve the descriptions of the analyses. Bunch suggested detailing sample size information, providing additional analysis on how many measures were collected, and providing more information on the ranges of estimated WTP values. Brownstone also noted that the database failed to detail the population used to construct the WTP estimates, for each study. He commented that this could be important information given that the populations used to estimate WTP may be very different, based on the methods used in the study. For example, stated preference studies might cover a more representative sample of the wider population, whereas revealed preference studies might only cover a subset of the population, such as those that made new vehicle purchases, or representatives from a specific geographic location.

Parsons provided additional comments on the description of studies and attributes included in the analysis. He thought the list of included studies and attributes were appropriate (including the box on Page 3-1). He was also impressed by the direct contact with study authors to uncover any additional studies that might not have been easy to locate through traditional literature search methods. He also felt that the tabulation by year and study type was helpful, but had an additional suggestion for the authors to provide an overlay of the 52 studies on Figure 3-1, in order to see the temporal aspects of the set of included studies.

Parsons also suggested a frequency distribution of the top 5 to 10 journals and their percent frequency in the sample used in the study. He also said that the authors could provide a unique

data set count similar to Table 3-1, but for datasets. This would be helpful as some tables used the same data, and it would provide further insight into how the data was used.

1.2 Are the Methods Described Sufficiently

Generally, the reviewers felt the methods were described in sufficient detail, though two of the reviewers gave suggestions for improvement.

Bunch commented on the WTP calculation methods. He noted that the study could have benefitted from a more formal structure, such as those used in meta-analysis studies, and more specifically meta-analyses of WTP data. Bunch felt the meta-analysis structure could add more specificity and rigor to the computation of WTP, and to measures of uncertainty both within individual studies and across studies.

Brownstone also commented that the authors were not sufficiently clear about how they addressed issues of heterogeneity when it exists in the target population. He noted that the authors should strive to estimate WTP for the target population on average and also a confidence band around this estimated WTP for each included study. He noted that this would avoid the current problem of the authors using heterogeneity that arises from the use of random parameters models to represent the uncertainty of the estimated WTP values.

Parsons recommended expanding the discussion of "other models with random distributions" on Page 4-1. He noted that these models are not described in the report. He also asked about the treatment of nested multinomial logit models, and if, for these models, it is possible to have more than one estimated WTP value for an attribute. He also asked if these attributes are included in more than one nest with different coefficients or scale estimates. Parsons also commented on how price variable might have entered into a WTP function non-linearly. Lastly, on Page 4-3, he noted that the authors should provide further detail on whether mean values were used for triangular, uniform, and other types of distributions that might have shown up in the included set of studies.

1.3 Are the Methods and Procedures Technically Appropriate

Overall, the reviewers found the methods and procedures to be technically appropriate. All three reviewers gave suggestions for improvement and refinement.

Bunch commented on the use and application of the Delta method. The issues he raised stemmed from the computation of statistical measures. The ratio of two normally distributed random variables results in a statistic having an undefined mean and variance. He noted that the statistical properties of quantities computed in this manner are often problematic, unless the denominator of the ratio is far from zero and has a very high t statistic. He added that the applications he has seen in the literature, where the Delta method has worked well, have involved denominators that were estimated with a high degree of precision.

Bunch recommended that the authors review a 2013 working paper by Carson and Czajkowski entitled "A New Baseline Model for Estimating Willingness to Pay from Discrete Choice Models," which discusses the issues with the Delta method in more detail. Carson and Czajkowski's example is based on simulation and shows the problems with using the Delta method when coefficients in the ratio are estimated with only moderate precision. As a partial remedy, Carson

and Czajkowiski suggest avoiding mean and variance statistics, and they recommend use of the mean in the calculations. Bunch commented that the issues described by Carson and Czajkowiski suggest why EPA might have found such a wide variance in the resulting WTP estimates.

To avoid the issue of a wide variance in WTP estimates, Bunch recommended using the median as the measure of central tendency, and to use alternate approaches for estimating confidence intervals. In addition, Bunch again suggested that the authors follow the approaches outlined in the meta-analysis literature, and in particular the approach followed by Van Houtven in a 2008 paper entitled "Methods for Meta-Analysis of Willingness to Pay Data: An Overview." He also suggested reviewing meta-analyses that has been conducted for estimating the value of travel time savings to see how meta-analysis approaches could be applied in this study.

Brownstone gave specific suggestions on how to calculate confidence bands around the WTP value estimates. He noted that the authors did not use formulas presented for measuring the bias and variance of the WTP estimates (equations C1 and C5) when calculating the confidence bands, which resulted in standard errors that are always too small.

Brownstone commented that a better approach would be to calculate the confidence bands using equations C1 and C5 to generate valid bounds for the bias and variance of WTP estimates. Brownstone gave a detailed example to note how this approach results in more realistic estimates of standard errors. He recommended that the estimates of confidence intervals for the WTP estimates be redone using his suggested approach.

Parsons also asked about the reasonableness of the procedures used by EPA, even though he found no specific technical errors. He suggested using simulation methods to calculate unbiased means, particularly given that EPA is already using simulation methods to estimate a dispersion of WTP estimates from mixed logit models.

Parsons also recommended that EPA should focus on the data from stated preference studies as opposed to the data from revealed preference studies. He noted that models based on revealed preference data are more likely to sketch out the marginal cost of attributes and not the demand function. Parsons illustrated this claim through a description of an estimation of the WTP for changes in the shoulder room attribute. He noted that with the data from revealed preference studies, what is estimated is the marginal WTP at the equilibrium implicit price for different consumer groups, but not estimates of the entire demand function. He believed that this is a concern due to the main policy interest of EPA of changes in WTP for vehicle attributes that are non-marginal in nature. Parsons felt that even though WTP data from stated preference studies is hypothetical, the nature of SP data is much better suited to estimating demand functions, and thus for WTP for non-marginal change in vehicle attributes.

1.4 Are the Results Presented Appropriately

The reviewers all agreed that the results are presented appropriately and are generally expected given the common findings in the literature. For example, EPA's findings of the wide variability of WTP results as well as the finding that some attributes have signs opposite of what would be predicted by economic theory have been observed commonly in other studies. The reviewers noted that some of these findings may thus be unavoidable and are an accurate representation of the challenges of estimating WTP for vehicle attributes. Despite this general

agreement on the findings and how EPA presented them, the reviewers did all express some specific concerns.

Bunch again stressed the benefits that a meta-analysis approach could bring to the study, such as alternative approaches for estimating WTP, and for assigning weights to the individual studies in the sample. He also noted that WTP values having the wrong sign can relate to estimates measured with low precision, and thus high standard errors. This issue can be particularly prevalent in certain population subgroups, such as high income populations. He again stressed the advantages of the approaches outlined by Carson and Czajkowski as a means for estimating the central tendency and spread of WTP estimates. The procedures they suggest could also be used to develop weights for individual studies.

Bunch also recommended looking directly at the tradeoff between the vehicle attributes of interest rather than evaluating the WTP for these attributes. He suggested that EPA should directly calculate the marginal rate of substitution between attributes of interest, such as fuel efficiency and performance, which might increase the precision of the resulting estimates due to some of the commonly noted issues that arise in the estimation of WTP. He noted, however, that there have also been concerns stated in the literature regarding the stability of the coefficients estimated for performance measures, as compared to other vehicle attributes. He stressed that he still recommends that EPA estimate and present WTP estimates, but that marginal rates of substitution between attributes may be a useful way of evaluating the tradeoffs that customers are willing to make between vehicle attributes of interest.

Bunch also stated that the WTP discussion in Section 4 omits some discussion and background around the theory of conditional indirect utility functions. He felt that the study could benefit from discussion of this type, such as the theoretical discussion included in by McFadden in papers from 1981 and 1998, entitled "Structural Analysis of Discrete Data and Econometric Applications" and "Measuring Willingness-to-Pay for Transportation Improvements," respectively. He also recommended including information on the computation of closely-related measures of consumer welfare, using log-sums from mixed and nested logit models.

Bunch also recommended that EPA should request that researchers compute WTP in their study results. He noted that if EPA is going to make this request, that they could also expand this request to include estimates of WTP based on log-sum measures. Bunch added that it would be helpful for researchers to estimate these measures because it can be difficult for others to estimate these measures without access to the original data.

Brownstone commented that variability in WTP estimates is a common finding in the literature, and agreed with EPA's suggestion that the estimates variability is too large to find useful values to serve as consensus of estimates for WTP. He reiterated the role that he believes the small confidence bands estimated by EPA played in this overall study finding. Brownstone noted that EPA should suggest the estimation of models in WTP space as advocated by Train et al. in a 2008 study entitled "Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps." Brownstone felt that the approach suggested by Train et al. greatly simplifies the calculation of confidence bands for WTP estimates, and makes it easier to model heterogeneity while still imposing restrictions on the signs of WTP estimates (which could help to alleviate the issues of WTP values having theoretically incorrect signs).

Overall, Parsons felt that Section 5 was logical and clear. However, he had two suggestions for this section. First, he noted that a discussion of vehicle class was missing. He also thought that this section could benefit from a more consistent use of graphs and tables across attributes.

Parsons also had concerns about the interpretation of the coefficients for the fuel type and vehicle class attributes. He noted that these WTP estimates originate from estimates of the coefficients of a group specific constant, which means that they must have been computed relative to some excluded group. He believed that, for fuel type, this excluded group must be conventional gas vehicles, even though this is not clearly stated in the report. His concern was that, because many of the key attributes that would distinguish gas from electric vehicles are already included in the models, the coefficients on the group specific constants are mostly picking up residual unobserved differences. He noted that, if all of the attributes that matter to consumers are already included in the models, then the coefficients for the group specific constants should be close to zero. Parsons mentioned this issue because he did not think the report text adequately demonstrated the meaning of these coefficients. For example, the report text read as if these constants are estimating WTP for the full difference between gas and electric vehicles, and not just the unobserved effect.

1.5 Additional Comments

The reviewers had some additional comments that did not fit within the specific headings of the charge questions and were thus provided separate from the charge question responses. For the most part, however, these additional comments reiterated issues that were already brought up by the reviewers in the charge question responses.

Bunch repeated his suggestion of reworking the approach for estimating WTP from individual studies using meta-analysis methods and approaches. One advantage that this approach would offer is that it would allow for an evaluation of the merits and sources of individual WTP measures at a very early stage of the study selection process. It would also provide a basis for developing quantitative weights for individual studies.

Brownstone reiterated a concern that many of the studies included were not really designed to measure WTP for vehicle attributes, but to predict demand for an exogenously determined set of real and hypothetical vehicles. He suggested a more appropriate approach for policy research on demand for vehicle attributes, which would involve forecasting based on scenarios involving choices among actual vehicles. For example, vehicle choice models could be used to discover the price subsidies needed to achieve a target market share for a realistic set of more efficient vehicles. A similar approach could also be used to estimate the potential welfare losses associated with changes to vehicle attributes. He also reiterated the suggestion to work in WTP space as suggested by Train et al. in their 2008 study. He noted that this adjustment to the approach would make it easier to compute consistent confidence intervals to represent uncertainty in the WTP measurements.

Brownstone also gave some specific suggestions for the study presentation. He noted that Figure 5-1 is hard to read in black and white. He recommended that, in this figure, each study should be a vertical line from top to bottom of the confidence interval, with a bold dot on the point estimate (similar to what is done in Figure 6-1 on Page 6-6). He also suggested that the lines be arrayed in order from low to high point estimates. He cautioned against the claim in

Figure 5-13, that air conditioning is an inconsequential vehicle attribute. He felt this assumption was not correct for studies looking at battery electric vehicles, due to how the use of air conditioning degrades vehicle range. Additionally, Brownstone noted that the claim, on the top of Page 5-14, about the difference between the stated and revealed WTP estimates does not seem to be supported by the values presented in Table 5-2. He also had similar concerns about the second to last paragraph on Page 6.1.

Parsons gave overall suggestions for simplifying the study, noting that it is currently too busy. He felt it includes too many aspects to be as useable as it could be in policy contexts. He suggested that EPA should narrow the focus to a key set of attributes that EPA could use in policy evaluations, such as fuel savings and performance measures. Narrowing the scope of the study would make it easier to digest, and would emphasize the policy considerations that are of most interest to EPA.

Parsons also suggested that EPA should modify the current approach of using means for individual characteristics from the sample where each study was done, to compute WTP by instead creating some representative consumer types and estimating WTP values for each of these from the chosen study. He recommended some possible consumer groups for this approach, such as those based on income, family size, or location.

Parsons also mentioned that EPA only used one WTP outcome per paper, which would be like having a preferred model. He noted that having many WTP estimates from one paper, or having one paper dominate the full set of studies, is misleading. Another approach would be to consider only one model per dataset. Parsons also reiterated his suggestion that EPA should focus on stated preference rather than revealed preference studies. He favors stated preference data due to fact that stated preference data is more suited than revealed preference data to estimating the demand curve for non-marginal changes in vehicle attributes.

Parsons also suggested that EPA conduct some directed simulations using a handful of studies. Selection of studies for these simulations could be based on the quality of theory utilized, methods, data used, and the capacity to perform realistic analyses. Having selected studies based on these characteristics, EPA could then estimate welfare effects for non-marginal changes in several attributes simultaneously using this simulation-based approach. The result would be realistic and policy-relevant scenarios designed around EPA's specific concerns. Parsons also suggested that EPA possibly conduct its own stated preference studies to look at the specific policy issues of their concern, but he understood why it may not be possible for EPA to conduct primary research in this area.

2. Comments by Reviewer

2.1 Comments by Dr. David Bunch

CHARGE QUESTION	COMMENTS	AUTHOR RESPONSES
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	**With the above discussion as background, I would first comment that, although the authors provide documentation on their process for identifying and collecting studies, I have concluded that there is not enough detail on the reasons for why some of the selection criteria were chosen. I review again here that, based on my (admittedly very quick) review of some of the meta-analysis literature, I think all of the steps performed in this work could have benefited from more structure and rigor of the type I saw described in that literature.** **In terms of specifics: The studies are limited to those involving US-based populations. The argument here is (presumably) that, because the policies are for the US market, there would be no value in including studies from, e.g., Europe. On the surface, this seems understandable and perhaps even reasonable. However, given the well-known concerns about the variability of WTP measures and the potential impact of sample sizes, research methodologies, etc., the existence of such studies could provide additional statistical power for evaluating these concerns, which I think are paramount. On the other hand (and in fairness to the authors), I suppose it could be argued that the extent**	

	of these issues might not have been fully appreciated ex ante, and that the extent of these concerns might not have been fully evident except as an outcome of this work. Having said this, this actually reinforces the importance of treating this aspect of the design process in a more rigorous way. **Similarly, a decision was made to use a cutoff of 1995. The rationale provided is not compelling, particularly if this cutoff were to eliminate studies that might have disproportionate value due to superior sample size, potential study quality, etc., given the well-known concern about potential variation in WTP estimates. For example, there are some relatively important studies that appeared in the small number of years prior to 1995 that could have been important to include.** **More generally, as I reviewed this work it seemed to me that there were studies that I am personally aware off that were probably not included, but under the circumstances it was difficult to sort this out (and I believe it would have been inappropriately time consuming). These studies could have been omitted for various reasons stated by the authors, but there was no efficient way for me to know which studies were eliminated, and why.** **In summary (and to directly address this specific charge question, as stated): I found that the current version of this review makes it rather difficult to evaluate the "quantity and quality of the data used for analysis." This response can also be linked back to some the introductory remarks provided above. The issue of sample size, in particular, has not been**	

	addressed in this study. More generally, the authors have collected a lot of measurements here, and it seems that it would be possible to do a more systematic exploration of a variety of important issues that have been identified, e.g., the issue of whether the type of data (e.g., RP, SP, market data, etc.), amount and structure of data (intertemporal and/or geographic variation), and modeling methodology systematically affects these estimates, and, if so, in what way. This could lead to more carefully considered conclusions on ranges for WTP estimates. (See later comments in response to other questions.)	
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	**The short answer to this question is a lightly qualified "yes." Based on my exploration of the material in Section 4 and Appendix C, I have some potentially serious concerns about how the calculations were performed in this study. I will go into greater technical detail in my response to the next question.** **However, as I general matter, I would echo my earlier remarks that my review of various literature references has led me to conclude that the design and execution of this study could have benefited from application of the more formal structure that has been developed in the meta-analysis literature in general, and on the topic of WTP in particular. This was already mentioned in the response to charge question 1 regarding the selection process for studies. In the next charge question, I will go into more detail on the technical question of computing WTP measures. But, more generally, the overall process should pay attention to issues**	

	related to computing the ES measures, and how to compute measures of uncertainty both within and across studies. This can culminate in development of more systematic models of these measures that shed light on the nature and source of these variations. Again, although I realize that this specific study clearly intended to "stop short" of performing so-called "meta-analysis regression modeling," I have concluded the procedures for properly performing the initial steps could have greatly benefited from adhering more closely to the meta-analysis paradigm that exists in the literature.	
3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	**A critical aspect of this study as that the vast majority of WTP measures are computed as ratios of estimated coefficients, and the statistical framework for producing the results is based on the Delta method as discussed in Section 4 and Appendix C.** **Although it seems to be not very well appreciated in much of the social science literature, in the statistics literature there are well-known problems that arise when a statistical measure is computed as the ratio of two normally distributed random variables: the distribution of this ratio is Cauchy, with undefined mean and undefined variance.** **In the report, the authors compute WTP measures using a ratio of estimated coefficients from discrete choice models in the vast majority of cases. The coefficient estimates have typically been obtained by maximum likelihood, and are treated as being asymptotically normal with some variance-covariance matrix. In various**	

	subfields of the social science literature, researchers have commonly used the Delta method (or a variant) for estimating statistics related to the WTP distribution. Another method includes parametric bootstrapping (Krinsky and Robb 1986). **However, statistical properties of quantities computed using these methods can be very badly behaved unless the estimated coefficient for the denominator variable (i.e., the price coefficient) is "statistically far" from zero. One result from the literature suggests the following rule of thumb: To use the Delta method, the t-statistic for the denominator coefficient should be above 8.75. (Note that this depends on the combined values of the coefficient estimate and the estimated standard error.) Now, there has been a variety of Monte Carlo studies in the social science literature that suggest the Delta method "works well." However, in the fine print it can be discovered that they frequently generate results using cases where the denominator coefficient has been estimated with a high degree of precision (as suggested by the rule of thumb mentioned above).** **Before continuing, one specific recommendation to the authors is that they obtain and study a working paper by Carson and Czajkowski (2013) ["A New Baseline Model for Estimating Willingness to Pay from Discrete Choice Models"] where the technical details mentioned above are discussed in more detail.** **In particular, they offer a simulation example showing that, when the coefficients estimates have been obtained**	

	with only moderate precision, that there are specific problems with the behavior of both the Delta and Krinsky-Robb methods. In particular, the mean and variance statistics should be avoided (as might be expected given that the statistic is, after all, Cauchy distributed). The median is the appropriate measure of central tendency, the 96% confidence interval/quantile range contains negative numbers, and it is skewed. At the same time, the Delta method yields a reasonable mean (even though it shouldn't), and a standard deviation that is not only finite (which is wrong) but also overly small, leading to confidence intervals that are both inappropriately narrow and symmetric (also wrong). **Taking a step back, and repeating from before: Recall that a major outcome of this report is the observation that the computed WTP results seem to "vary widely," and that this raises a variety of concerns. In addition, the authors also note that, simultaneously, the computed estimates of "precision" for WTP measures within a given study are rather high, which seems to be a contradiction. However, note that this outcome would appear to be entirely consistent with the discussion given above regarding what can happen when using the Delta method in this situation. Specifically, Carson and Czajkowski (2013) explains in detail why the procedure followed by the authors is ill-advised.** **Have gone over this, I would add that I don't have the wherewithal to give the authors step-by-step instructions on**	

	exactly how to do this instead. I want to be clear that I also understand appreciate that the authors are facing constraints in being forced to use only the information that is available from existing published studies. However, I do have some ideas about possible directions. At the very least, it is clear that the measures of central tendency should probably be the median and not the mean to the degree this is possible, and some alternative approach to computing a confidence interval (or perhaps, e.g., an interquartile range) should be used for measuring uncertainty (i.e., not the standard deviation). The material in section III of Carson and Czajkowski (2013) would seem to provide some possible avenues for doing this. Having said this, we must acknowledge that almost any procedure developed on this basis could run into the same difficulties the authors faced in this study, i.e., that certain key statistics (such as covariance estimates) might not be available. **More generally, the authors should take a more careful look at the literature on WTP statistical measures, and also at actual meta-analyses of WTPs that appear in the literature. [However, with regard to the statistics, my view is that the Carson and Czajkowski (2013) paper should be the touchstone: There are some published results in the transportation literature that one should be careful about when considering, e.g., there is a paper by Gatta, et al. (2015) in Transportation Research A on computing WTP confidence intervals that would seem to make some misleading claims about some results reported in Daly, Hess, and deJong (2012), who**	

	recommend using the Delta method. Carson and Czajkowski (2013) do also cite the Daly et al. paper, but only as an example of a paper that recommends using the Delta method. It is clear from their paper that they disagree with and reject this conclusion.] One reference that specifically addresses meta-analysis of WTP is by George Van Houtven (2008) ["Methods for the Meta-Analysis of Willingness-to-Pay Data: An Overview," Pharmacoeconomics 2008: 26 (11): 901-910]. **One point of contrast we can make here is that there are meta-analyses in the transport literature for another important transport-related WTP measure: Value of Travel Time Savings. I have concluded that this would seem to be a much "easier" problem than WTP for vehicle attributes, and it may be that WTP measures are measured with so much more precision that the types of problems described above do not occur. Examples are by Wardman and co-authors, e.g., Wardman et al. (2016) Transportation Research A. It would seem to be the case model estimates yield more precise estimates (in contrast to vehicle choice models), for example, outcomes that are unstable to changes in specification and/or having the incorrect sign do not seem to be much of an issue. In fact, in this literature the analyses seem to be routinely performed using ln(WTP) as the dependent variable, which does take into account skewness. At the same time, they seem to use mean values (without difficulty) rather than try to develop a more complex procedure to ascertain medians. Moreover (on a**	

	different point emphasized above), there does seem to be attention paid to addressing variation due to differences in sample size, and there are various technical details regarding whether to treat variation due to differences in studies as fixed effects versus random effects. Moreover, there is attention paid to how results should be analyzed when multiple measures are used from the same study. (Note that I have added these remarks for the sake of completeness, to further reinforce my recommendation that the authors pay more attention to research design issues related to meta-analysis.) **To conclude, the material provided above in response to this charge question is primarily focused on technical issues around computing WTP-related statistics. However, it is also clear that these issues are linked to the broader context of the study itself, as addressed in the previous paragraph. This provides the basis for moving on to the next charge question.**	
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	**This question focuses on presentation of results, and conclusions. To review a few points from my introductory comments and the response to previous questions, the structure and content of this report can be summarized now as follows:** **1. Studies were identified and selected. WTP measures were computed, largely along the lines as described in Section 4 and Appendix C.** **2. Results on WTP measures and confidence intervals were summarized from individual studies in Appendix B. Summary measures across studies were**	

	presented in various forms (graphical and tabular) in Section 5. **3. Basic conclusions about the results include the major observation that these WTP distributions have wide variation, and frequently have the theoretically incorrect sign. At the same time, the within-study estimates of variation seem to frequently suggest more precision.** **4. The apparent "lack of consensus" in these WTP results is discussed and addressed in Section 6. As part of this discussion, the authors begin to explore in more detail the possible reasons for why these values might vary across studies, and start to go into more detail by looking at more specific subsets of results from different studies.** **At this stage, I have a couple of inter-related concerns about the status of this work. The first concern would probably be obvious from comments I provided to charge question 3. It is unclear to what degree some of the issues with the results are a function of the specific details of how the WTP distributions were computed and summarized. At the same time, the technical background I provided also sheds some light on why these measures may have potentially unavoidable difficulties due to the challenges associated with model estimation.** **Having said this, the technical discussion also suggests that there may be better ways to quantitatively evaluate the relative merits of results coming from different studies at an earlier stage of the process, by adhering to various procedures**	

	researchers have developed for performing meta-analysis. Specifically, in meta-analysis it would appear that measures from different studies are routinely given WEIGHTS that capture at least some of these concerns. **Now, let's digress for a moment regarding some of the current results. First, there is the problem that sometimes WTP measures have "the wrong sign." There may be no way to finesse this when faced with the authors' challenge of using existing results. But, consider the fact that estimates with the wrong sign might also frequently be measured with low precision (i.e., large standard errors). Price coefficients in vehicle choice models are notoriously difficult to get precise estimates for, and as show in some of the results, when researchers attempt to estimate price coefficients for income-based segments, incorrect signs can emerge (typically for higher income households). It may be that some judgment is required when evaluating this situation, keeping in mind that researchers were addressing multiple issues and were not solely focused on producing WTP estimates for policy analysis purposes. A similar situation would apply if a price coefficient is relatively small (but with the correct sign), leading to a very high WTP estimate. Even if the estimate is statistically significant, unless it is measured with an adequate level of precision, this will yield all of the statistical problems described in the previous section.**	

	In other words, it seems to me that a more quantitative way of discriminating across studies (e.g., using weights) should be developed, and they should be used much earlier in the process when producing results of the type that are currently presented in section 5. (I realize that at this stage the message might be repetitive: the meta-analysis paradigm and other examples in the literature might provide a useful guide on how to do this.) **At this stage, I need to try to "draw a line" and move on to concluding remarks. One issue here is that I want to be constructive as a reviewer, and I have tried to keep reminding myself that some of the thoughts I am generating here might be viewed as a function of "20-20 hindsight."** **To summarize in a more focused way: The technical issues raised in response to charge question 3 should probably be considered carefully by the authors. It may be that it is possible to use some of the theory in Carson and Czajkowski (2013) to come up with an alternative way of computing a measure of WTP central tendency and precision based on quantiles (where central tendency is the median) for the individual observations, and, if possible, that these also should be used for the purpose of developing weights. Having suggested this, it will almost certainly be the case that some type of approximation or working assumption will be required similar to the one currently made for implementing the Delta method. However, this could end up producing better results, and potentially a contribution to the literature. Next, finding**	

	a way to address the relative precision of measures from various studies and using appropriately weighted results rather than treating all results as equal would be a potential improvement. It might be possible to make these comparisons using summary statistics. **This is where I recognize that I am getting even closer to the "slippery slope" of suggesting that, once the above things have been addressed, it should be straightforward to do a meta-analysis regression that would yield quantitative results to address some of the issues currently discussed in Section 6.** **One final thought worth considering: we have already mentioned the difficulty of getting precision on price coefficient estimates. However, one of the primary issues of concern right now is the tradeoff in consumer preferences involving fuel efficiency versus performance. Would it make sense to do some type of exploration of how consumers directly trade off these attributes versus each other? In other words, rather than use intermediate estimates of WTP, directly compute marginal rates of substitution for these two attributes. It seems likely that the coefficients might have better precision. On the other hand: our experience is that coefficients for performance measures tend to be more unstable than for other attributes. One additional idea here is that this approach would necessarily involve a specific set of studies to compute these tradeoffs, and the differences could be compared across studies. The current situation presents results for these**	

	attributes separately, and drawing conclusions that are based on comparing the behavior of WTP estimates for each of the two attributes could be compromised by the fact that different studies are used for each separate analysis. **The current "conclusions" section includes a variety of pieces of advice for researchers going forward. Some of this is okay to say, but the authors know it would also be difficult. Researchers have lamented the lack of, e.g., "validation studies," for a very long time but there have always been practical obstacles. (However, as data availability increases this might be more reasonable.)** **The authors suggest that researchers should also routinely produce WTP estimates from their models and report them to avoid the issues faced by the authors in compiling this report. Fair enough.** **BUT: If you are going to go down this route, you might want to include additional details that would be helpful in this regard, given the serious technical issues associated with using ratios of estimates as is done here. Specifically, note that Carson and Czajkowski (2013) directly address this type of issue, giving a suggestion for how coefficients should be estimated to avoid some of the issues associated with using ratios of coefficient estimates. However, note that their suggestion is tantamount to enforcing a strictly negative price coefficient (in a way that there is no statistical mass at zero). This gets into the potentially dicey territory of how much a researcher should impose**	

	theory-based restrictions on estimates. (This is rather subtle, since in this case the approach is more targeted toward the statistical variation in the estimate than the estimate itself.) **Second, I have avoided until now pointing out that the WTP discussion in Section 4 leaves out the actual underlying theory based on the conditional indirect utility function. It could be important to include some of this information in a report like this one, which can be traced back to McFadden (1981). More specific treatments relevant to this paper are, e.g., McFadden (1998) ["Measuring Willingness-to-Pay for Transportation Improvements," in Theoretical Foundations of Travel Choice Modeling, Garling, Laitila, Westin eds., Pergamon], although there are other similar papers that could be looked at.** **In particular, some of the well-known results for computing closely-related measures of consumer welfare using log-sums from MNL and NL models could be worth considering. As long as researchers are being asked to compute WTP measures, why not ask them to use log-sum-based measures if they happen to be using MNL or NL? One question I have not had the band-width to explore is under what conditions these measures end up being equivalent to the coefficient ratio measure used in this paper. In any case, it is certainly likely that these measures would not have the same statistical problems as the ratio of coefficients. These are, of course, the types of measures that simply cannot be easily computed except by the researchers at the**	

	time they perform a study, because they require full access to all of the original data.	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):		
At this stage, I think the approach I have taken has allowed me to cover a variety of issues in conjunction with the charge questions themselves. I continue to encourage the authors to consider finding a way to incorporate more of the principles of the "meta-analysis paradigm" in conjunction with addressing the more specific items identified above. One major implication is that the issue of evaluating the relative merits of the sources of individual WTP measures should be more formally addressed at a much earlier stage, and based on some quantitative criterion that can also be used as "weights." **[Bunch has also posted the following comments outside of the original table. ICF has copied in here.]** **The following material constitutes my review and comments. The primary structure followed is as prescribed in the instructions we were given in the charge letter, i.e., to generate responses to specific Charge Questions in tabular form. To improve the efficacy of the process of generating these responses as well as readability, I have taken the liberty of slightly revising the tabular format.** **In addition, I would add the following: There are aspects of the review request discussed in the charge latter that appear in the paragraphs preceding the charge questions themselves. I consider them to be 1) important and relevant, but 2) not necessarily adequately reflected in the Charge Questions themselves. Although it might have been possible to postpone this discussion to the last sections of the table, I decided that I wanted to instead include some introductory material prior to answering the specific questions. Some of the issues were also explored in more depth during our kickoff conference call, where we posed a number of questions and discussed some preliminary reactions to the report.** **First, here are some comments in response to specific excerpts from the charge letter.** **"EPA has been developing estimates of the willingness to pay for (WTP) a variety of vehicle characteristics, such as fuel economy, performance, and comfort. If vehicle standards that EPA issues affects these other vehicle characteristics, having estimated values for these attributes might enable EPA to develop better estimates of the benefits and costs of its standards. … The estimates are based on existing published studies of**		

light-duty vehicle demand in which authors have econometrically estimated purchase patterns for vehicles based, in part, on vehicle characteristics." **Before responding to the charge questions more directly, it will be helpful to review at a very high level what this report consists of. The work basically consists of:** **1. A systematic literature review consisting of identifying and collecting studies that perform data analyses and/or model estimation yielding results that can be used to compute WTP estimates for various vehicle attributes.** **2. An explanation of methods used for computing WTP estimates.** **3. Summarization of computed WTP statistics and distributions for various individual attributes, at both the level of individual studies, and compiled across studies.** **4. Discussion related to what was observed in the results. This was primarily focused on the fact that WTP estimates appear to exhibit large variability across studies. Some of this discussion involved considering in more detail specific results from selected studies to explore possible reasons for this variability, attempting to evaluate the possible impact of various research design dimensions (e.g., different data types, model types, estimation approaches, etc.).** **Again, the key overall finding (that WTP estimates from the vehicle choice modeling literature appear to vary widely, and even take on theoretically incorrect signs) is unsurprising and has been an ongoing concern of researchers and policy makers alike for quite some time. We discussed this problem/issue during the kickoff call, because it would seem to beg the question of whether or not this type of project should not have been more concerned with trying to evaluate the relative quality of the studies, rather than simply collecting and reporting numbers. To clarify, the charge letter itself does include the following:** **"We ask that you review methods and underlying assumptions, their consistency with the current science as you understand it, and the clarity and completeness of the presentation. For this review, no independent data analysis is required. Rather, we ask that you assess whether the data and methods are applied appropriately, given the state of current understanding, and the conclusions reasonably drawn."** **We asked for, and received clarification that this part of the request primarily applies to the "methods and underlying assumptions" of the review itself, not the underlying studies on which the review is based. However, what appears next is:** **"Note that EPA's interest in these estimates is based not only in seeking robust estimates for values of willingness to pay for vehicle characteristics, but also in understanding the robustness of the models from which these estimates are derived."**	

Note that this expands the scope of EPA's interest to understanding "robustness of the models from which these estimates are derived," in order to add clarity to the relative robustness of the estimates themselves. This was a subjected we attempted to explore during the phone call. During this discussion, we suggested/asked whether or not this study shouldn't go more in the direction of "meta-analysis," rather than stop short by limiting the review to the collection of data and reporting of summary statistics. In either case, it seems that finding a way to evaluate the relative merits of various "sources" of "WTP estimates" is potentially important to EPA. **These are the issues that I found myself concerned about finding a way to address, in an attempt to provide something of value to EPA in my review. To this end, I spent some time researching the subject of meta-analysis (which has not been a specific area of specialty for me). One thing I concluded through this effort is that it would be incorrect to view "doing a meta-analysis" as an extension of the steps already taken in this review.** **Rather, I concluded that the steps performed for this review should actually be viewed as corresponding to the initial steps within the overall process of "doing a meta-analysis." In particular, designing the procedures for identifying and collecting studies, creating rules for determining which studies to include or exclude, specifying the technical details for computing measurements and relevant statistics for "Effect Sizes" (or, "ESs," which in this case are WTPs for attributes), deciding on a rule for eliminating outliers, and even producing summaries of Effect Size distributions represent the steps of a meta-analysis that then typically culminate in building models for the ES data that seek to identify and clarify the specific sources of "variation" across studies.** **To be clear, there is usually a process that also weights the relative importance of data points from different studies using the inverse of a measure of "internal variance" for the study, based on whatever the most appropriate statistics are for the situation being studied. These variance measures are typically a function of sample size, e.g., studies with small sample size are expected to yield more biased and/or noisy estimates and therefore should be "down-weighted." Moreover, these weights are also frequently employed when producing the reported summary statistics (such as those appearing in this review). All of these steps, taken together, are important aspects of designing a meta-analysis that addresses the needs and goals of the research. For example, even the first step of defining how the universe of studies is determined should be based on these considerations.** **I wanted to report this, since at the time of our phone call I tried to articulate a general concern about whether or not relative "study quality" should somehow be addressed in light of EPA's interests, and, if so, how that might be achieved. I believe this additional background on meta-analysis sheds some light on this question. The above remarks**	

represent an overall sentiment that I will keep coming back to in the more detailed comments that follow.	
ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:	
Given the specific direction that this review took, I have elected to forgo more detailed and specific items in individual sections.	

2.2 Comments by Dr. David Brownstone

CHARGE QUESTION	COMMENTS	AUTHOR RESPONSES
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	**Almost – they appear to have done an excellent job finding all of the relevant studies – at least I am not aware of any that should be included. However the key database they construct (Appendix B) fails to note the population used to construct the WTP estimates for each study. For example some studies use market data and therefore include all light vehicle sales – including rental cars, commercial fleets, etc. Other studies use survey data that may only cover new personal vehicle purchasers or possibly a representative sample of some geographic area. If there is heterogeneity in WTP across the US population, then different WTP estimates could be partially explained by different underlying populations.**	
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	**The report is clear about the methods used to calculate WTP and the "confidence bands" around these estimates, but they are not clear about how to handle models that include heterogeneity across the target population. Ideally they should compute the population average WTP and a confidence band around this quantity for each study. While this may not be possible for some studies, it would be better if this was clearly elucidated as the goal of the exercise. This would make population heterogeneity a secondary issue, but currently the report sometimes uses heterogeneity arising out of random parameters to represent estimation uncertainty.**	

3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	**The report does an excellent job (in Appendix C) of describing the issues with getting unbiased estimates of WTP from standard choice models. They also give a reasonable formulas for the bias and variance of this WTP measure in equations C-1 and C-5. Unfortunately they then ignore these formulas when computing their "confidence bands" as described in the middle of page 4-4. The "standard errors" they calculate are always too small.** **A better approach would be to calculate the bands using equations C-1 and C-5 assuming the bounds given by -σα σβ ≤σαβ≤ σα σβ. The endpoints of this interval can be estimated and put into formulas C-1 and C-5 to generate valid bounds for the bias and variance of the WTP estimates. For example in a recent study I calculated WTP = 4 with a correct standard error of .09. The bias in this WTP estimate using equation C-1 lies between 0 and .02. The approach used in the report gives a standard error estimate of .012 (a very big underestimate), while the using the bounds on σαβ in equation C-5 imply that the correct standard error lies between 0 and 0.4. Unless the bias in the WTP estimates using the above bands are small, then the tables and figures should be redone showing the bands around the WTP estimates. Likewise the "confidence intervals" should be redone using the bounds.**	
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	**The overall conclusion is that there is too much variability in the existing estimates to find any useful "consensus" values. Since the "confidence bands" computed in the report are always too small (see 3.**	

	above) the key conclusion follows from the results. The report offers some reasons for the lack of accurate WTP estimates, and these are generally reasonable. Since the purpose is to estimate WTP I am surprised that they did mention estimating the models in WTP space as advocated by Kenneth Train (see reference below). This greatly simplifies computing confidence bands for the WTP estimates and also makes it easier to model heterogeneity in WTP while still enforcing sign restrictions.	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):		
This report is based on an exhaustive review of the recent US-based literature on the willingness to pay (WTP) for different vehicle attributes. These WTP measures are typically compared to cost estimates to provide an additional unit of the attribute, but it is not clear how this corresponds to behavior in the real vehicle marketplace. Indeed many of the papers reviewed by this report were not designed to "measure" WTP, but instead to predict demand for an exogenously-specified set of real and hypothetical vehicles. Near-term policy evaluation might be done better using scenario forecasts with realistic vehicles. For example instead of looking at the difference between WTP and the engineering cost of a 1c/mile reduction in operating costs, the vehicle choice models could be used to discover the price subsidies needed to achieve a target market share for a realistic set of more efficient vehicles. Alternatively the welfare loss of achieving a particular fuel economy goal could be measured. I will take the goal of measuring WTP as given for the rest of this review. If authors of the papers reviewed were asked to estimate WTP now, I suspect that many of them would reformulate their models to work in WTP space (Kenneth Train, Riccardo Scarpa and Mara Thiene, Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps, American Journal of Agricultural Economics, Vol. 90, No. 4, pp. 994-1010, 2008.). This would allow for directly modeling population heterogeneity in WTP. More importantly it would make it much easier to compute consistent confidence intervals to represent the uncertainty in the WTP measurements.		

ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:	
Figure 5.1 These figures are hard to read - especially in black and white. I suggest representing each study by a vertical line extending from the top to the bottom of the “Confidence Interval” with a bold dot on the point estimate (similar to Figure 6-1 on page 6-6 in the report). These lines could be arrayed from low to high point estimates as in the current graphs. **Section 5.1.3 - the claim that air conditioning is “perhaps inconsequential” since it is almost standard equipment is not correct for studies looking at battery electric vehicles. The use of air conditioning seriously degrades vehicle range for these vehicles, leading to a potential demand for new technologies like better vehicle thermal insulation and (possible solar-powered) vehicle ventilators.** **Page 5-14 - the claim on the top about the differences between stated and revealed WTP estimates does not seem to be supported by the values in Table 5.2. See also second to the last paragraph on page 6-1.**	

2.3 Comments by Dr. George Parsons

CHARGE QUESTION	COMMENTS	AUTHOR RESPONSES
1. Does the presentation describe the choice of publications used for the estimates sufficiently to allow the reader to form a general view of the quantity and quality of data used for the analysis?	**The Literature Review (Section 2) and Description of Studies and Attributes (Section 3) are clear, complete, and systematic. The coverage of studies is good and the method used to identify and include studies is good. The box on page 3-1 and supporting text are a good description of the selection process. The studies include the best current and recent-past research as far as I could tell. The direct contact with authors, which apparently uncovered additional studies, is impressive and makes this more thorough than most investigations of this sort. The tabulation by year and study type is also very helpful in understanding the data. I have a few suggestions:** **(1) It would be useful to see an overlay of the 52 "main sample" studies on figure 3-1 to see where the final selections fall over time.** **(2) It would be useful to see a frequency distribution for the top 5 (maybe 10) journals represented in the sample and their percent frequency in the sample.** **(3) In addition to a paper count, it would be interesting to see a unique data set count, since some of the 52 studies use the same data. Sort of a Table 3-1 for data sets.** **(4) Somewhere early in the study, we should be told more about how EPA will use the numbers – maybe an exploratory example application. I kept wondering how will these many different points estimates**	

	would be used to get at welfare effects. I know EPA does an effort like this for VSL's (i.e., summarizing literature to get at a VSL to use in policy analysis). In that case, we know you will be multiplying the VSL by number of lives saved or lost. It is easier. But here, the welfare effects come through market choices in response to a policy (fuel standards). It is more difficult. Having some idea upfront of how the numbers will be used, would help. I might be missing something obvious here! **Overall, the presented description is sufficient to form a general view of the quality and quantity of the data used.**	
2. Does the presentation describe the methods sufficiently to allow the reader to form a general view of the quality and validity of the calculations used in the development of the willingness-to-pay (WTP) estimates?	**The Methodology (Section 4) is mostly clear and complete. The description of how the "central tendencies of WTP" are calculated is straightforward and the discussion on why the first order approximations are used is good. I have several suggestions here (I number these beginning where I left off above for easy reference):** **(5) On page 4-1 (point 3 near the top), there is reference to "...other models with random distributions…" What other models are these? They are never described or discussed as far as I could tell.** **(6) On page 4-3 (second full paragraph), for NMNL models you write "..the utility functions of the nests that include the prices of vehicles are used in estimating WTP.." Is it possible to have more than one value for an attribute from a given study because price-attribute pairs show up in more than one nest with different**	

	coefficient estimates or different scale estimates? If so, how was this handled? **(7) You do not mention cases where price might enter non-linearly (e.g., ln(Pr)). Are there any of these and, if so, how are they handled?** **(8) On page 4-3, you write that they use "..mean values for normally distributed random coefficients and median values for lognormally distributed coefficients..". I assume you use means for triangular, uniform and other distributions that show up in the literature, but nothing was stated, so we cannot be sure.** **Aside from these points, the description of the methods is sufficient to form a general view of the quality and quantity of the WTP calculations.**	
3. Are the methods and procedures employed technically appropriate and reasonable? In areas where RTI, based on data limitations, has made assumptions to conduct the calculations, are the assumptions appropriate and reasonable? Please distinguish between cases involving reasonable disagreement in methods as opposed to cases where you conclude that current methods involve specific technical errors.	**I found no technical errors but do have a couple questions about reasonableness of the procedures.** **(9) As I read the manuscript, you use simulation to estimate a dispersion of WTP estimates from mixed logit models but do not use it for computing means. Why not simulate to calculate unbiased means?** **(10) Given that the automobile market is highly competitive, I believe the models based on RP data largely sketch out marginal cost, not demand functions. Think about variation in the price of a car due to shoulder room. Adding x square feet of shoulder space comes at an added cost of Δc to the manufacturer. Competitive forces, in theory, should force the price difference for a small should-room car versus a large shoulder-room car**	

	(all else constant) to be near Δc, so Δp=Δc -- a standard result from micro-theory. This should hold for all attributes over all levels. The price differences, in principle, then capture marginal cost differences in RP data. (The classic identification problem in the hedonic setting is that RP data sketch out an implicit price schedule and not demand or supply – true, but in a case of competitive market like autos, that price schedule will be a cost function.) As far as WTP goes, you will have an estimate of marginal WTP at the equilibrium implicit prices for different consumer groups, but not the demand function. I think this is important to keep in mind when analyzing the results. Since the primary interest is WTP and ultimately non-marginal changes in WTP, since changes in fuel standards and any ensuing changes in comfort, performance, etc., are likely to be non-marginal, you need to be careful. For this reason, I give the nod to SP data in this setting -- despite its being hypothetical. The price changes paired with the varying attribute levels in the SP case are not market-based. Instead, they are set by researchers, and should in principle be sketching out demand functions and hence WTP for non-marginal changes. This comment falls in the "reasonable disagreement" category. **Otherwise, I thought the procedures were appropriate and reasonable.**	
4. Has RTI presented the results of the analysis in appropriate ways? Do the conclusions follow logically from the results?	**The results are presented mostly in an appropriate and reasonable way. Section 5 was logical and clear. I noticed the discussion of Vehicle Class was missing. Also, the section could perhaps benefit**	

	from a more consistent use of graphs and tables across attributes, but that is not major. I suppose some deserve more attention than others. I do have a concern about how Fuel Type and Vehicle Class are interpreted. **(11) The WTPs for Fuel Type and Vehicle Class come from Group Specific Constant coefficient estimates, which must be relative to some excluded group. For Fuel Type I am assuming the excluded group is a conventional gasoline vehicle. It is never clearly stated or discussed in the text, but gas is excluded in Table 5-1. Since many of the key attributes that distinguish gas from electric and other vehicles are included in the models, the Group Specific Constants are picking up average "unobserved" differences between gas and electric (or whatever other fuel type is in question). Of course, most of what is driving the difference in gas versus electric is likely to be included in the variables like range, fuel time, performance, etc., so the Group Specific Constant WTPs are picking up the residual differences only. Indeed, if you included all attributes that matter to people, you would expect these constants to be near zero. In any case, the text does not seem appreciative of meaning of these coefficients and their WTPs. I may be misreading the text but they read more like they are interpreting the WTP as picking up the full difference between a gas and electric vehicles and not a residual unobserved effect. Some refinement is needed here. Vehicle Class has the same issue. In this case, I am also wondering how they kept this exclude group**	

	(whatever it is) constant across the studies.	
ADDITIONAL OVERALL COMMENTS PROVIDED (NOT CHARGE QUESTION-SPECIFIC):		
I have a few suggestions for simplifying the analysis and directing it more toward its final use in policy. As it stands now, it is just "too busy" and needs some researcher judgement calls to make it more manageable and useful for policy. I know the wide range of values makes application seem almost futile but EPA needs to proceed somehow and I think the literature has value. Here are my suggestions: **(12) Focus on the key set of attributes that EPA will use for policy analysis. Clearly fuel savings and some of the performance measures stay, but narrowing the set would make it easier to digest and bring focus on to what matters. Do we really need to look at air conditioning, fuel availability, auto-transmission, and so on?** **(13) Instead of using means for individual characteristics from the sample where each study was done to compute WTP, consider creating some representative consumer types and estimate values for each of these from each chosen study. The groups might be defined by income, family size, and location.** **(14) Consider using one WTP outcome per paper, a sort of preferred model. Having many estimates from one paper or one paper dominating the set of studies for one attribute is misleading. Just because an author of a study considers another model, they get greater representation. You could even consider one model per data set.** **(15) Focus on SP studies for reasons I mentioned above.** **Finally, in a somewhat different but related vein, I also think it is worthwhile to consider some "directed" simulations using just a handful of studies:** **(16) These would be selected based on the quality of theory, method, and data used and capacity to perform realistic analysis (i.e., include a useful array of attributes). Then, consider estimating welfare effects for non-marginal changes in several attributes simultaneously using these model. These would be realistic, policy-relevant scenarios meaning to capture EPA concerns (e.g., fuel economy improves, performance drops, and comfort drops). These might use a log-sum calculation, in which case the analyst would have to somehow formulate a representative choice set. Honestly, I have not thought through the details and may be missing potential pitfalls, but I like the idea of treating some studies as better than others and choosing them for application and getting values of direct use for policy. This approach also allows for substitution, which is likely to be critical for getting at the true welfare effects – something separate point estimates of price do not. Another strategy is to do your own primary analysis using SP data and**		

targeting the specific policy questions at hand. You could even do some contingent behavior analyses to get at specific issues of policy relevance. This is probably you first best solution.	
ADDITIONAL COMMENTS BY SPECIFIC REPORT CHAPTER:	

Appendix A. Resumes of Selected Reviewers

1. Dr. David Bunch

DAVID S. BUNCH

Curriculum Vitae
March 2016

Current address:
Graduate School of Management
One Shields Avenue
University of California, Davis
Davis, CA 95616
USA

Phone: 530-752-2248
Fax: 530-752-2924
e-mail: dsbunch@ucdavis.edu

Education

Ph. D., Rice University, 1985 (Mathematical Sciences)
Master in Applied Mathematical Sciences, Rice University, 1981
M. S., Northwestern University, 1979 (Chemistry)
B. A. (cum laude), Rice University, 1978 (Chemistry)

Positions

Professor of Management, UC Davis, July 2000-present
Acting Director, 'Center for New Mobility Studies,' Institute of Transportation Studies, UC Davis, October 1999-August 2000.
Associate Professor of Management, UC Davis, July 1992-July 2000.
Visiting Scholar, Department of Marketing, Faculty of Economics, University of Sydney. July 1997-July 1998.
Assistant Professor of Management, UC Davis. July 1985-June 1992.
Visiting Assistant Professor, UC Davis. July 1984-June 1985.
Associate, Rice Center, Houston, Texas. May 1982-August 1983.
Research Associate, The Institute for Rehabilitation and Research, Houston, Texas. February 1980-January 1982.

Courses taught

Product Management
Marketing for E-Commerce
Marketing Research
Management Policy
Decision Making and Management Science
Marketing Models for New Products
Discrete Choice Analysis
Managerial Decision Making
Systems Analysis and Design

Applied Linear Models for Management
Special Topics in Management of Information Systems
Seminar in Management

Publications and Papers

"A Theory-Based Framework for Incorporating Behavioral Effects from Discrete Choice Models into Bottom-Up Energy Sector Models," in preparation for submission to *Energy Economics*.

"Variance-Component-Based Nested Logit Specifications: Improved Formulation, and Practical Microsimulation of Random Disturbance Terms," with David M. Rocke, *The Journal of Choice Modelling*, 21 (2016) 30–35).

"Incorporating Behavioral Effects from Vehicle Choice Models into Bottom-Up Energy Sector Models," with Kalai Ramea, Sonia Yeh, and Christopher Yang, Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-15-13 (2015).

"Fuel Economy and CO2 Emissions: Standards, Manufacturer Pricing Strategies, and Feebates," with C. Liu and D. L. Greene, The Energy Journal, Vol. 35, No. 3 (2014), pp. 71-89.

"Numerical Methods for Optimization-based Model Estimation and Inference," Handbook of Choice Modelling, S. Hess and A. Daly, eds., Edward Elgar, Cheltenham UK (2014), Chapter 23, pp. 565-598.

"Recent Advances on Modeling Multiple Discrete-Continuous Choices," with C. Bhat and A. Pinjari, Choice Modelling: The State of the Art and the State of Practice, S. Hess and A. Daly, eds., Edward Elgar, Cheltenham UK (2013), Chapter 3, pp. 73-90.

"Impacts of Feebates in Combination with Fuel Economy and Emissions Standards on U.S. Light-Duty Vehicle Fuel Use and Greenhouse Gas Emissions," with C. Liu, D. L. Greene, E. C. Cook, Transportation Research Board, Paper 11-2027.

"Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California" (with D. L. Greene, T. Lipman, E. Martin, S. Shaheen), California Air Resources Board, Final Report on Contract UCD 08-312, University of California, Davis, CA, February, 2011.

"Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California: Interim Statement of Research Findings" (with David L. Greene). Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-10-13

"Follow-on Development of CARBITS: A Response Model for the California Passenger Vehicle Market," Final Report (Contract 05-303) prepared for State of California Air Resources Board, April 30, 2009.

"Exploring the Consumer Valuation of Organic-Related Properties in Fresh Produce Choice," (with Yuko Onozaka and Doug Larsen), Working Paper.

"Theory-based Functional Forms for Analysis of Dissagregated Scanner Panel Data" Working Paper.

"Behavioral Models and Estimates for Leisure-Passenger Value of Travel-Time Saving in Long-Hall Air Travel Markets Using Stated Choice Experiments" Working Paper.

"Behavioral Frontiers in Choice Modeling" (with Wiktor Adamowicz, Trudy Ann Cameron, Benedict G. C. Dellaert, Michael Hanneman, Michael Keane, Jordan Louviere, Robert Meyer, Thomas Steenburgh and Joffre Swait), Marketing Letters, Volume 19, Numbers 3-4 (December, 2008), pp. 215-228.

"Automobile Demand and Type Choice," Handbook of Transport I: Transport Modeling (Second Edition), with Belinda Chen, in David A. Hensher and Kenneth J. Button, editors, Pergamon, (2008) pp. 463-479.

"Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles" (with David Brownstone and Kenneth Train) in The Stated Preference Approach to Environmental Valuation, Volume III: Applications: Benefit-Cost Analysis and Natural Resource Damage Assessment, ed. Richard T. Carson, Series: The International Library of Environmental Economics and Policy, Ashgate Publishing (2007).

"Hybrid Choice Models: Progress and Challenges," (with Moshe Ben-Akiva, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, Axel Boersch-Supan, David Brownstone, Andrew Daly, Andre de Palma, Dinesh Gopinath, Anders Karlstrom and Marcela A. Munizaga), Chapter 11 in Transportation Planning, Classics in Planning Series, Volume 7, eds. Yoram Shiftan, Kenneth Button, Peter Nijkamp, Edward Elgar Publishing (May, 2007).

"California Air Resources Board – Institute of Transportation Studies (CARBITS) Vehicle Market Microsimulation Model for California, UC Davis Institute of Transportation Studies Research Report (2004).

"Purchase Pattern Analysis for Organic and Conventional Produce with Detailed Supermarket Scanner Data," Working Paper, with Doug Larson and Yuko Onozaka (2004).

"Hybrid Choice Models: Progress and Challenges," (with Moshe Ben-Akiva, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, Axel Borsch-Supan, David Brownstone, Andrew Daly, Andre de Palma, Dinesh Gopinath, Anders Karlstrom, Marcela A. Munizaga), Marketing Letters, 13(3): pp. 163-175 (August 2002).

"Optimal Designs for 2^k Paired Comparison Experiments," (with Deborah J. Street and Beverly Moore), Communications in Statistics - Theory and Methods, 30(10), pp. 2149-2171 (2001).

"Automobile Demand and Type Choice," Handbook of Transport I: Transport Modeling, David A. Hensher and Kenneth J. Button, editors, Pergamon (2000), pp. 463-479.

"The First-Passage Approach to Valuing the American Put," (with Herb Johnson). Journal of Finance, 55 (5), (October 2000), pp. 2333-2356.

"Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles" (with David Brownstone and Kenneth Train). Transportation Research B, Volume 34, Issue 5 (June 2000), pp. 315-449.

"Combining Sources of Preference Data for Modeling Complex Decision Processes" (with Jordan J. Louviere, Robert J. Meyer, Richard Carson, Benedict Delleart, W. Michael Hanemann, David Hensher, and Julie Irwin). Marketing Letters, Volume 10, Issue 3 (August 1999), pp. 205-217.

"Discussion of 'Multi-Featured Products and Services: Analyzing Pricing and Bundling Strategies' by Moshe Ben-Akiva and Shari Gershenfeld" (with Jordan J. Louviere), Journal of Forecasting, Volume 17, Issue 3-4 (1998), pp. 197-201.

"Determinants of Alternative Fuel Vehicle Choice in the Continental United States" (with Melanie Tompkins, Danilo Santini, Mark Bradley, Anant Vyas, and David Poyer), Transportation Research Record, Number 1641, Energy and Environment: Energy Air Quality, and Fuels 1998, Transportation Research Board, National Research Council.

"Commercial Fleet Demand for Alternative-fuel Vehicles," (with Thomas F. Golob, Jane Torous, David Brownstone, Soheila Crane, and Mark Bradley), Transportation Research A Vol. 31A (1997): 219-233.

"A Vehicle Usage Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilization Data," (with Thomas F. Golob and David Brownstone), Journal of Transport Economics and Policy Vol. 31 (1997): 69-92.

"Analysis of the Future Household Market for Alternative Fuel Vehicles in Southern California Using a Microsimulation Forecasting System," (with Camilla Kazimi), University of California-Davis, Graduate School of Management Working Paper UCD GSM WP# 05-96 (1996).

"A Comparison of Experimental Design Strategies for Choice-Based Conjoint Analysis with Generic-Attribute Multinomial Logit Models (with Jordan J. Louviere, and Don Anderson). (1996, revised version of UCD GSM Working Paper 01-94).

"A Dynamic Forecasting System for Vehicle Markets with Clean-Fuel Vehicles," (with David Brownstone and Thomas F. Golob). In D. A. Hensher, J. King, and T. H Oum eds., World Transport Research, Volume 1 (1996): 189-203.

"A Vehicle Transactions Choice Model for Use in Forecasting Demand for Alternative-Fuel Vehicles," (with David Brownstone, Thomas F. Golob, and Weiping Ren), Research in Transportation Economics, Vol. 4 (1996): 87-129.

"A Demand Forecasting System for Clean-Fuel Vehicles," (with David Brownstone and Thomas F. Golob), in Organization for Economic Co-operation and Development (OECD) Towards Clean Transportation: Fuel Efficient and Clean Motor Vehicles, Publications Service, OECD, Paris, France, 1996, 609-624.

"Experimental Analysis of Choice," (with J. Louviere, R. Carson, D. Anderson, P. Arabie, D. Hensher, R. Johnson, W. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J. Wiley), Marketing Letters, Vol 5:4, 351-368.

"Extension of the Four-Parameter Logistic Model for ELISA to Multianalyte Analysis," (with G. Jones M. Wortberg, S. Kreissig, S. Gee, B. Hammock, and D. Rocke), Journal of Immunological Methods, Vol. 177, 1-7.

"Demand for Clean-Fuel Vehicles in California: A Discrete-Choice Stated Preference Survey" (with Mark Bradley, Thomas F. Golob, Ryuichi Kitamura, Gareth P. Occhiuzzo). Transportation Research A, Vol. 27A, No. 3, pp. 237-253, 1993.

"Predicting the Market Penetration of Electric and Clean-fuel Vehicles" (with Thomas F. Golob, Ryuichi Kitamura, and Mark Bradley), The Science of the Total Environment, 134 (1993) pp. 371-381.

"Subroutines for Maximum Likelihood and Quasi-Likelihood Estimation of Parameters in Nonlinear Regression Models" (with David M. Gay, and Roy E. Welsch). ACM Transactions on Mathematical Software, Vol. 19, No. 1, March 1993, Pages 109-130.

"Who Deters Entry? Evidence on the Use of Strategic Entry Deterrents" (with Robert Smiley). Review of Economics and Statistics, Vol. 74, No. 3 (August 1992), pp. 509-521.

"A Simple and Numerically Efficient Valuation Method for American Puts Using a Modified Geske-Johnson Approach" (with Herb Johnson). Journal of Finance, Vol. XLVII, No. 2 (June 1992), pp. 809-816.

"Estimability in the Multinomial Probit Model," Transportation Research B, 1991, Vol 25B(1), pp. 1-12.

"Statistical Design of ELISA Protocols," Journal of Immunological Methods, 1990, (with David Rocke and Robert Harrison), 132 (1990), 247-254.

"Heterogeneity and State Dependence in Household Car Ownership: A Panel Analysis Using Ordered-Response Probit Models with Error Components," 11th International Symposium on Transportation and Traffic Theory, Elsevier, July 1990 (with Ryuichi Kitamura).

"Multinomial Probit Model Estimation Revisited: Testing Estimable Model Specifications, Maximum Likelihood Algorithms, and Probit Integral Approximations for Trinomial Models of Household Car Ownership," Transportation Research Group Research Report UCD-TRG-RR-4, April 1990, (with Ryuichi Kitamura).

"How Many Choices Are Enough?' The Effect of the Number of Observations on Maximum Likelihood Estimator Performance in the Analysis of Discrete Choice Repeated-Measures Data Sets with the Multinomial Logit Model," UC Davis Graduate School of Management Working Paper UCD-GSM-WP0390, January 24, 1990, (with Richard Batsell).

A Comparison of Algorithms for Maximum Likelihood Estimation of Finite Mixture Density Models. Working paper.

"When Are Additive Models Valid for Evaluating Proposed Research?," Methods of Information in Medicine, 1989, 28, pp. 168-177 (with D. Cardus, M. J. Furher, R. M. Thrall).

"A Monte Carlo Comparison of Estimators for the Multinomial Logit Model" (with Richard R. Batsell), Journal of Marketing Research, February 1989, pp. 56-68.

"A Comparison of Algorithms for Maximum Likelihood Estimation of Choice Models," Journal of Econometrics, May/June 1988, Vol. 38, No. 1/2, pp. 145-167.

"Efficient Algorithms for Maximum Likelihood Estimation of Probabilistic Choice Models, SIAM Journal of Scientific and Statistical Computing, 1987, 8(1), 56-70.

"Parameter Estimation of Probabilistic Choice Models," Department of Mathematical Sciences Technical Report TR84-5 (Ph.D. Thesis), Rice University, 1984.

Presentations [Not necessarily complete.]

"Incorporating Behavioral Effects from Vehicle Choice Models into Bottom-Up Energy Sector Models," (work with Kalai Ramea, Sonia Yeh, and Chris Yang), Sustainable Mobility Seminar (Winter 2016), Stanford University, January 22, 2016.

"Incorporating Behavioral Effects from Vehicle Choice Models into Bottom-Up Energy Sector Models," (work with Kalai Ramea, Sonia Yeh, and Chris Yang), Transportation Seminar Series, University of Tennessee Department of Civil & Environmental Engineering, October 28, 2016.

"Incorporating Behavioral Effects from Vehicle Choice Models into Bottom-Up Energy Sector Models," (work with Kalai Ramea, Sonia Yeh, and Chris Yang), International Conference on Operations Research 2015 (OR2015), University of Vienna, September 4, 2015.

"Aggregation Bias in Discrete Choice Models with an Application to Household Vehicle Choice," (work with Timothy Wong and David Brownstone), 14th International Conference on Travel Behavior Research (IATBR), Windsor, England, July 20, 2015.

"Aggregation Bias in Discrete Choice Models with an Application to Household Vehicle Choice," (work with Timothy Wong and David Brownstone), presented at the International Conference on Choice Modelling, University of Texas, Austin, May, 2015.

"Aggregation Bias in Discrete Choice Models with an Application to Household Vehicle Choice," (work with Timothy Wong and David Brownstone), presented (by David Brownstone) to Centre for European Economic Research (ZEW), Manheim, Germany, October 2014.

"Making MARKAL-TIMES Models Behave As Though They Include (Nonlinear) Discrete Choice Models," (work with Kalai Ramea, Sonia Yeh, and Chris Yang), U.S. DOE Vehicle Choice Modeling Workshop, University of California, Davis, October 1, 2014.

"Review of (the Frequently Forgotten) Microeconomic Theory That Underlies Discrete Choice Models," NSF-Sponsored Workshop on Complex Choice Behavior and Transportation Energy Policy, Ann Arbor, Michigan, September 29, 2014.

"Aggregation Bias in Discrete Choice Models with an Application to Household Vehicle Choice," (work with Timothy Wong and David Brownstone, presented by David Brownstone), presented to Danish Technical University Transportation Group, Kongens Lyngby, Denmark, September 2014.

"Aggregation Bias in Discrete Choice Models with an Application to Household Vehicle Choice," (work with Timothy Wong and David Brownstone), presented (by David Brownstone) at the Stockholm Fuel Price Conference in Stockholm, Sweden, September 2014.

"Modifying Optimization-based Energy Systems Models to Incorporate Discrete Choice Models of Vehicle Choice and Usage Behavior," (work with Kalai Ramea, Sonia Yeh, and Chris Yang), presented at LOT 2014 Conference, Molde, Norway, September 2, 2014.

"Estimating Demand for High Fuel Economy Vehicles," (work with David Brownstone) presented at the International Conference on Choice Modeling (ICMC) in Sydney, Australia, July 2013.

"Incorporation of Consumer Demand in Energy Systems Models and their Implications for Climate Policy Analysis," (work with Kalai Ramea, Sonia Yeh, Chris Yang, and Joan Ogden), presented (by Kalai Ramea) to 32nd International Energy Workshop, Paris, France June 19-21, 2013.

"Energy Commission Models for Analyzing Household-Based Transportation Energy Demand: Evaluation, Model Improvement Options, and Recommendations," (with David Brownstone), Final Project Presentation, California Energy Commission, June 30, 2014.

"Incorporation of Consumer Demand in Energy Systems Models and their Implications for Climate Policy Analysis," (work with Kalai Ramea, Sonia Yeh, Chris Yang, and Joan Ogden), presentation to California Energy Commission, February 28, 2013.

"Demand for High Fuel Economy Vehicles," (work with David Brownstone), presented at the 13th International Conference on Travel Behavior Research (IATBR), Toronto, Canada, July, 2012.

"Recent Advances in Modeling Multiple Discrete-Continuous Choices," Invited Workshop Presentation, Second International Choice Modelling Conference, Oulton Hall, Leeds, July 5, 2011.

"Potential Impacts of Feebate Programs for New Passenger Vehicles," (with D. L. Greene and T. E. Lipman), California Air Resources Board – Cal/EPA HQ, Sierra Hearing Room, June 14, 2011.

"Economic Incentives for New Vehicle Purchases to Reduce Greenhouse Gas Emissions: Research on Policy Options for California," Transportation Center Seminar Series, Northwestern University Wednesday, August 12, 2009.

"University of California Feebates Research Project," Kitamura Memorial Symposium, June 29, 2009, University of California, Davis.

"University of California Feebates Research Project," presented at the Fuel Economy/Greenhouse Gas Emissions Standards – Technical Meeting, California Air Resources Board, June 17, 2009.

"Feebate Policy Workshop," (with David L. Greene and Tim Lipman) California Air Resources Board, February 26, 2009.

"Theory-based Functional Forms for Analysis of Dissagregated Scanner Panel Data," presented to workshop on Behavioral Frontiers in Choice Models, Seventh Tri-Annual Choice Symposium, Wharton, June 2007.

"Theory-based Functional Forms for Analysis of Dissagregated Scanner Panel Data," presented to the 2007 Bay Area Marketing Consortium, May 11, 2007.

"Recent Advances in Discrete Choice Models," (with Jordan J. Louviere), tutorial workshop, Thirteenth Annual Advanced Research Techniques (A/R/T) Forum, Vail, Colorado (June 2-5, 2002).

"Identifying Optimal Offerings and Campaigns In Interactive Channels Using Real-Time Experiments and Automated Modelling Procedures," (with Moshe Ben-Akiva, Denis Bolduc, Richard Carson, Jordan Louviere, Hikaru Phillips, Matthew Symons), Thirteenth Annual Advanced Research Techniques (A/R/T) Forum, Vail, Colorado (June 2-5, 2002).

"Information and Sample Size Requirements for Estimating Non-IID Discrete Choice Models Using Stated-Choice Experiments," 2001 UC Berkeley Invitational Choice Symposium, Asilomar Conference Center, Pacific Grove, California, June 1-5, 2001.

"Estimation of Non-IIA Discrete Choice Models," Winter Quarter Econometrics Seminar Series, Department of Economics, UC Davis (February 19, 1999).

"Implications of Choice Task Complexity Effects for Design and Analysis of Discrete Choice Experiments" (with Jeff D. Brazell and Jordan J. Louviere). Presented at 1998 INFORMS Marketing Science Conference, INSEAD, Paris (July 1998).

"Estimating Non-IIA Models Using Discrete Choice Stated-Preference Data: Model Forms, Sample Size Effects, and Simulation Estimation." Presentation to workshop on "Combining Sources of Preference Data for Modeling Complex Decision Processes," HEC Invitational Choice Symposium, Paris (July, 1998).

"Estimating Non-IIA Models Using Discrete Choice Stated-Preference Data: Model Forms, Sample Size Effects, and Simulation Estimation." Presentation to PhD seminar, Department of Marketing, University of Sydney, Australia (May 27, 1998).

"Optimal Designs: Discussion." Presented at "Workshop on Experimental Design and Experimental Data: Alternative Perspectives," Department of Econometrics, University of Sydney, Australia (May 8, 1998).

"Computational Methods for Maximum Likelihood Estimation." Presentations to Ph.D. course in "Bayesian Estimation Methods," AGSM (Australian Graduate School of Management), University of New South Wales, Sydney, Australia (May 1, May 15, 1998).

"Determinants of Alternative Fuel Vehicle Choice in the Continental United States" (with Tompkins, Santini, Bradley, Vyas, and Poyer), Presented at 1998 Transportation Research Board Meetings, Washington, D. C. (January 1998).

"Random Parameter Logit Models to Forecast Vehicle Ownership" (with David Brownstone and Kenneth Train), in preparation. Presented at the Eighth Meeting of the International Association of Travel Behavior Research, September 1997, University of Texas, Austin.

"Analysis and Forecasts of EV Markets: Background and Methods for Multi-Year Household Surveys." Presented at "Electric Vehicle Markets: Conceptual and Analytical Approaches for Understand EV Demand," Institute of Transportation Studies, UC Davis (November 20, 1996).

"Analysis and Forecasts of EV Markets: Results for Multi-Year Household Surveys." Presented at "Electric Vehicle Markets: Conceptual and Analytical Approaches for Understand EV Demand," Institute of Transportation Studies, UC Davis (November 20, 1996).

"Using Dynamic Microsimulation and an Integrated System of Revealed Preference and Stated Preference Discrete Choice Models to Estimate Penetration of Electric Cars in the California Vehicle Market" (with D. Brownstone and T. F. Golob), presented at the 1996 INFORMS Marketing Science Conference, March 1996, University of Florida, Gainesville.

"Testing a multinomial extension of partial profile choice experiments: Empirical comparisons to full profile choice experiments" (with Keith Chrzan and Daniel C. Lockhart), presented at the 1996 INFORMS Marketing Science Conference, March 1996, University of Florida, Gainesville.

"Forecasting future vehicle usage using a jointly-estimated revealed- and stated-preference model" (with T. F. Golob and D. Brownstone). Presented at the Annual Meeting of Transportation

Research Board, National Research Council, National Academy of Sciences, Washington, D.C., January 7-11, 1996.

"The future of alternative fuel vehicles in California: Projections from a microsimulation forecasting system" (with D. Brownstone and T. F. Golob). Presented at the Annual Meeting of Transportation Research Board, National Research Council, National Academy of Sciences, January 7-11, 1996, Washington, DC.

"Using Stated Preference and Intended Transactions to Predict Market Structure Changes for the Personal Vehicle Market in California," (with D. Brownstone, and T. F. Golob), presented at the 1995 Marketing Science Conference, July 1995, University of New South Wales, Sydney, Australia.

"Design strategies for experimental choice sets: Comparison of methods for multinomial logit models," (with J. Louviere and D. Anderson), presented at the 1994 Marketing Science Conference, March 1994, Univeristy of Arizona, Tuscon, AZ.

"A Demand Forecasting System for Clean-Fuel Vehicles," (D. Brownstone, D. S. Bunch, and T. F. Golob), presented at the OECD Conference "Fuel Efficient and Clean Motor Vehicles," Mexico City, March 28-30, 1994.

"Choice models from experimental choice sets," Workshop 2 Participant at Duke Invitational Symposium on Choice Modeling and Behavior (no-formal-presentation format), August 1993, Fuqua School of Business, Durham, North Carlolina.

"Predicting the market penetration of electric and clean-fuel vehicles," (T. F. Golob, R. Kitamura, M. Bradley, D. S. Bunch), presented at International Symposium on Transport and Air Pollution, September 10-13, 1991, Avignon, France.

"Modelling the choice of clean fuels and clean-fuel vehicles," (R. Kitamura, M. Bradley, D. S. Bunch, and T. F. Golob), presented at the PTRC Annual Meeting, September 9-12, 1991, University of Sussex, England.

"Demand for clean-fuel personal vehicles in California: A discrete-choice stated-preference survey," (with M. Bradley, T. F. Golob, R. Kitamura, and G. Occhiuzzo), presented at the Transportation Research Board Conference on Transportation and Global Climate Change: Long Run Options, August 25-28, 1991, Asilomar Conference Center, Pacific Grove, California.

"Advances in Computation, Statistical Methods and Testing," Workshop 1 participant, Banff Invitational Symposium on Consumer Decision Making and Choice Behavior (no-formal-presentation format), May 8-15, 1990, Banff, Alberta, Canada.

"Heterogeneity and State Dependence in Household Car Ownership: A Panel Analysis Using Ordered-Response Probit Models with Error Components," (with Ryuichi Kitamura) presented at TIMS/ORSA Joint National Meeting, Las Vegas, May 1990.

"Multinomial Probit Model Estimation Revisited: Testing Estimable Model Specifications, Maximum Likelihood Algorithms, and Probit Integral Approximations for Trinomial Models

of Household Car Ownership," (with Ryuichi Kitamura), presented at the 69th Annual Meeting of the Transportation Research Board, Washington, D. C., January, 1990.

"A Panel Analysis of Car Ownership Using the Multinomial Probit Model," (with Ryuichi Kitamura) Fall ORSA/TIMS meeting, Denver, October 1988.

"How Many Choices Are Enough? The Effect of Replications on MLE Performance in the Analysis of Discrete Choice Repeated-Measures Data Sets," invited presentation at the Joint Statistical Meetings of the American Statistical Association and the Biometric Society, August, 1988.

"A Monte Carlo Comparison of Estimators for the Multinomial Logit Model," presented at the Fall ORSA/TIMS meeting, St. Louis, October 1987.

"A Comparison of Algorithms for Maximum Likelihood Estimation of Choice Models," presented at the SIAM Conference on Optimization, Houston, May, 1987.

"Efficient Algorithms for Maximum Likelihood Estimation of Probabilistic Choice Models." Invited presentation for Computer Science and Statistics: the 18th Symposium on the Interface, Fort Collins, Colorado, March 1986.

Grants and Contracts

[Additional material to be added.]

University of California Center for Energy and Environmental Economics, 2011-2012. "The Demand for High Fuel Economy Vehicles," (with D. Brownstone).

California Air Resources Board, 2009-2011. Potential Design, Implementation, and Benefits of a Feebate Program for New Passenger Vehicles in California (with David L. Greene).

California Air Resources Board, 2005-2009. Follow-on Development of CARBITS: A Response Model for the California Passenger Vehicle Market.

California Air Resources Board, 2002-2004. Analysis of Auto Industry and Consumer Response to Regulations and Technological Change, and Customization of Consumer Response Models in Support of AB 1493 Rulemaking (with D. Sperling and A. Burke).

University of California Energy Institute, 1996-1997. An Evaluation of Policies Related to Vehicular Energy Use (with Golob and Brownstone).

California Energy Commission, 1995-1996. Development of Policy Sensitive Transportation Forecasting Models for Personal, Commercial Fleet, and Freight Activity (with Golob, Brownstone, and Kitamura).

Pacific Gas and Electric, 1993-1995. Alternative Vehicles in the Pacific Gas and Electric Service Area: A Project for Developing Models and Scenario Simulation Systems for Forecasting

AFV Penetration and Usage. (Principal investigator, with Co-PI's Golob, Kitamura, and Brownstone).

Southern California Edison, 1992-1994. Forecasting Electrical Vehicle Ownership and Use in the California South Coast Basin (with Golob, Kitamura, and Brownstone).

United States Department of Transportation, 1992-1993. Improved Designs for Stated Preference Analysis of Transport-Choice Processes (continuation).

United States Department of Transportation, 1991-1992. Improved Designs for Stated Preference Analysis of Transport-Choice Processes.

National Institutes of Health, National Institute of Environmental Health Sciences, 1988-1992. Decision Support System for Statistical Analysis of Toxics Measurement Data (with Rocke). Renewed for 1993-1996 (with Rocke).

California Energy Commission, 1991-1992. Assessing the Potential Acceptance of Alternative Fuels and Vehicles in California's Commercial Fleets (with Golob and Kitamura)

California Energy Commission, 1990-1992. Clean Vehicles/Clean Fuels Stated Preference Pilot Study (with Golob and Kitamura).

United States Department of Transportation, 1990-1991. Impact of Telecommuting on Travel: Accessibility Implications of Working at Home (with Kitamura, Jovanis, and Mokhtarian).

United States Department of Transportation, 1988-1990. Evaluation of the Impact of Telecommuting on Travel Patterns, Road Congestion, Energy Use and Air Quality (with Kitamura).

Professional Societies

INFORMS (Institute for Operations Research and Management Science)
Society for Industrial and Applied Mathematics

Editorial Board Journal of Choice Modeling

Reviewing

Reviewer for:
- ACM Transactions on Mathematical Software
- Annals of Operations Research
- Computational Statistics and Data Analysis
- Department of Transportation (UC Transportation Center)
- Energy Economics
- Global Environmental Change
- IEEE Transactions on Signal Processing
- Journal of the American Statistical Society
- Journal of Applied Econometrics
- Journal of Choice Modelling

Journal of Computational and Graphical Statistics
Journal of Business and Economic Statistics
Journal of Econometrics
Journal of Forecasting
Journal of Marketing Research
Mathematical Programming
Marketing Science
National Science Foundation
SIAM Journal on Optimization
SIAM Journal on Scientific and Statistical Computing
Transportation
Transportation Research
Transportation Research Board (TRB)
Transportation Science

2. Dr. David Brownstone

February 2016

DAVID BROWNSTONE, Professor of Economics

Department of Economics
University of California
3151 Social Science Plaza
Irvine, California 92697-5100

Telephone: (949) 824-6231
FAX: (949) 824-2182
EMAIL: dbrownst@uci.edu
WWW: http://www.economics.uci.edu/~dbrownst/

Education

A.B. University of California, Berkeley (1975) Applied Mathematics
M.A. University of California, Berkeley (1977) Statistics
Ph.D. University of California, Berkeley (1980) Economics

Academic Positions

1999- Professor of Economics, School of Social Sciences, University of California, Irvine

1992-99 Associate Professor of Economics, School of Social Sciences, University of California, Irvine

1984-92 Assistant Professor of Economics, School of Social Sciences, University of California, Irvine

1982-89 Senior Research Fellow, The Industrial Institute for Economic and Social Research, Stockholm, Sweden

1982-84 Assistant Professor, Department of Economics, Stockholm School of Economics, Stockholm, Sweden

1980-81 Visiting Assistant Professor, Department of Economics, Stockholm School of Economics, Stockholm, Sweden (Fall and Summer)

1979-82 Assistant Professor, Department of Economics, Princeton University, Princeton, New Jersey

Publications (Working paper versions in Adobe acrobat files are available if underlined, * denotes graduate student co-author)

(P37) "A Vehicle Ownership and Utilization Choice Model With Endogenous Residential Density" (with H. Fang*), Journal of Transportation and Land Use 7(2), 135-151, 2014.

(P36) "The Impact of Residential Density on Vehicle Usage and Fuel Consumption" (with J. Kim*), Energy Economics 40, 196-206, 2013.

(P35) "Methodological Developments in Activity-Travel Behavior Analysis," in R. M. Pendayala, C. R. Bhat (Eds.), Travel Behavior Research in an Evolving World. (pp. 249-260). International Association for Travel Behavior Research. ISBN 978-1-105-47378. 2012.

(P34) "The Impact of Residential Density on Vehicle Usage and Energy Consumption" (with T. Golob), Journal of Urban Economics, 65, pp. 91-98, 2009.

(P33) "Key Relationships Between the Built Environment and VMT." Paper commissioned by the National Research Council for Transportation Research Board Special Report 298:

Driving and the Built Environment, 2008. Accessible at: http://onlinepubs.trb.org/Onlinepubs/sr/sr298appendixb.pdf.

(P32) "Risk, uncertainty and discrete choice models" (with A. de Palma, M. Ben-Akiva, C. Holt, T. Magnac, D. McFadden, P. Moffatt, N. Picard, K. Train, P. Wakker and J. Walker), Marketing Letters, 19, pp. 269-285, 2008.

(P31) "Recent Progress on Endogeneity in Choice Modeling" (with J Louviere, K Train, M Ben-Akiva, C Bhat, T A Cameron, R . Carson, J.R. DeShazo, D Fiebig, W Greene, D Hensher, and D Waldman), Marketing Letters, 16 (3-4), 2005.

(P30) "Estimating Commuters' "Value of Time" with Noisy Data: a Multiple Imputation Approach" (with S. Steimetz*), Transportation Research B, 39, 865-889, 2005.

(P29) "Valuing Time and Reliability: Assessing the Evidence from Road Pricing Demonstrations" (with K. Small), Transportation Research A, 39, 279-293, 2005. French version published as: "Expériences de tarification routière en Californie: enseignements pour l'évaluation du temps et de la fiabilité," in André de Palma and Emile Quinet, eds., *La Tarification des Transports: Enjeux et défis* (*Transport Pricing: Stakes and Challenges*), Paris: Economica, 2005.

(P28) "Drivers' Willingness-to-Pay to Reduce Travel Time: Evidence from the San Diego I-15 Congestion Pricing Project." (with C. Kazimi*, A. Ghosh*, T.F. Golob, and D. van Amelsfort), Transportation Research A, 37, pp. 373-387, 2003.

(P27) "Hybrid Choice Models: Progress and Challenges," (with M. Ben-Akiva, D. McFadden, K. Train, J. Walker, C. Bhat, M. Bierlaire, D. Bolduc, A. Boersch-Supan, D. Bunch, A. Daly, A. de Palma, D. Gopinath, A. Karlstrom, M. A. Munizaga), Marketing Letters, 13:3, pp. 163-175, 2002.

(P26) "The Bootstrap and Multiple Imputations: Harnessing Increased Computing Power for Improved Statistical Tests." (with R. Valletta), Journal of Economic Perspectives, 15(4), pp. 129-141, Fall 2001.

(P25) "Discrete Choice Modeling for Transportation." in D. Hensher (ed.), Travel Behaviour Research: The Leading Edge. Amsterdam: Pergamon, pp. 97-124, 2001.

(P24) "Modeling Non-ignorable Attrition and Measurement Error in Panel Surveys: An Application to Travel Demand Modeling." (with C. Kazimi* and T.F. Golob), in Groves, R.M., D. Dillman, J.L. Eltinge and R.J.A. Little (eds.), Survey Nonresponse. New York: Wiley, pp. 373-388, 2001.

(P23) "Joint Mixed Logit Models of Stated and Revealed references for Alternative-fuel Vehicles." (with D. S. Bunch and K. Train), Transportation Research B, 34, 315-338, 2000.

(P22) "Bootstrap Confidence Bands For Shrinkage Estimators." (with Camilla Kazimi*), Journal of Econometrics, 90, pp. 99-127, 1999.

(P21) "Forecasting New Product Penetration With Flexible Substitution Patterns." (with K. Train) Journal of Econometrics, 89, pp. 109-129, 1998.

(P20) "Multiple Imputation Methodology For Missing Data, Non-Random Response, And Panel Attrition." In T. Gärling, T. Laitila and K. Westin, eds., Theoretical Foundations of Travel Choice Modeling. Pp. 421-450, Elsevier, 1998.

(P19) "A Vehicle Use Forecasting Model Based on Revealed and Stated Vehicle Type Choice and Utilisation Data," (with T. Golob and D. Bunch), Journal of Transport Economics and Policy, 31:1, pp. 69-92, 1997.

(P18) "Multiply-Imputed Sampling Weights for Consistent Inference with Panel Attrition," (with Xuehao Chu*) in Golob, T.F., R. Kitamura and L. Long, eds. Panels for Transportation Planning: Methods and Applications, pp. 259-276, Boston: Kluwer Academic Publishers, 1997.

(P17) "A Demand Forecasting System for Clean-Fuel Vehicles," (with D. Bunch and T. Golob), in Organization for Economic Co-operation and Development (OECD) Towards Clean Transportation: Fuel Efficient and Clean Motor Vehicles, Publications Service, OECD, Paris, France, 1996, pp 609-624.

(P16) "A Transactions Choice Model For Forecasting Demand For Alternative-Fuel Vehicles." (with D.S. Bunch, T.F. Golob and W. Ren*) in S. McMullen (ed.) Research in Transportation Economics, Volume 4, pp 87-129, JAI Press Inc., 1996.

(P15) "A Dynamic Forecasting System For Vehicle Markets With Clean-Fuel Vehicles." (with D.S. Bunch, T.F. Golob) in D.A. Hensher, J. King and T.H. Oum, eds., World Transport Research: Proceedings of 7th World Conference on Transport Research, 1: 189-203. Pergamon, Oxford, 1996.

(P14) "Commercial Fleet Demand for Alternative-fuel Vehicles in California," (with T. Golob, J. Torous, S. Crane, D. Bunch, and M. Bradley), Transportation Research A, 31:3, pp. 219-233, 1997.

(P13) "Modeling Earnings Measurement Error: A Multiple Imputations Approach," (with R. Valletta), The Review of Economics and Statistics, 78:4, pp. 705-717, 1996.

(P12) "Transportation Energy Use," (with Charles Lave) in R. Gilbert (ed.) The Environment of Oil, Kluwer Academic Publishers, Boston, 1993, pp. 11-42.

(P11) "The Effectiveness of Ridesharing Incentives: Discrete-choice Models of Commuting in Southern California," (with Thomas Golob), Regional Science and Urban Economics, Vol. 22, 1992, pp. 5-24.

(P10) "Bootstrapping Admissible Linear Model Selection Procedures," in R. LePage and L. Billard (eds.), Exploring The Limits of Bootstrap, Wiley, New York, 1992, pp. 327-344.

(P9) "Zoning, Returns to Scale, and the Value of Undeveloped Land," (with A. DeVany), Review of Economics and Statistics, Vol 73, 1991, pp. 699-704.

(P8) "The Demand for Housing in Sweden: Equilibrium Choice of Tenure and Type of Dwelling," (with P. Englund), Journal of Urban Economics, Vol. 29, 1991, pp. 267-281.

(P7) "Bootstrapping Improved Estimators for Linear Regression Models," Journal of Econometrics, Vol. 44, 1990, pp. 171-187.

(P6) "Efficient Estimation of Nested Logit Models" (with K. Small), Journal of Business and Economic Statistics, Vol. 7, 1989, pp. 67-74.

(P5) "Tax Reform and Housing Demand: The Distribution of Welfare Gains and Losses" (with P. Englund and M. Persson), European Economic Review, Vol. 32, 1988, pp. 819-840.

(P4) "A Microsimulation Model of the Swedish Housing Market" (with P. Englund and M. Persson), Journal of Urban Economics, Vol. 23, 1988, pp. 179-198.

(P3) "Effects of Tax Reform on the Demand for Owner-Occupied Housing: A Microsimulation Approach" (with P. Englund and M. Persson), Scandinavian Journal of Economics, Vol. 87 (4), 1985, pp. 625-646.

(P2) "Microeconometrics," in G. Eliasson (ed.), Microeconometrics: IUI Yearbook 1982-1983, Industrial Institute for Economic and Social Research, Stockholm, 1983, pp. 81-90.

(P1) "Econometric Models of Choice and Utilization of Energy-Using Durables," in J. Hausman (ed.), Choice and Utilization of Energy-Using Durables, Electric Power Research Institute, Palo Alto, California, 1980, Chapter 5.

Reports

Brownstone, D., Madanat, S., Hansen, M. (2010). Review of "Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study. Senate Transportation Committee, State of California.

Brownstone, D., Bunch, D. (2013). A Household New Vehicle Purchase Model to Support Analysis of the Impact of CAFÉ Standards. Brookings Institution.

Bunch, David S. and David Brownstone (2014). Energy Commission Models for Analyzing and Projecting Household Transportation Energy Demand: Evaluation, Model Improvement Options, and Recommendations. California Energy Commission.

Recent Research Grants

4/16 – 3/17 UCCONNECT (University of California Transportation Center), $172,314, Co-Principal Investigator with Michael McBride

8/15 – 7/16 UCCONNECT (University of California Transportation Center), $66,982

1/15 – 3/16 UCCONNECT (University of California Transportation Center), $119,594, Co-Principal Investigator with Michael McBride

6/13 – 6/14 California Energy Commission (subcontract through UC Davis), $33,750

1/12 – 6/15 UC Sustainable Transportation MRPI, $162,575

10/11 – 7/13 Brookings Institution, $40,000.

9/10 – 6/12 University of California Energy and Environment Institute, $33,878.

1/09 – 8/10 U.C. Davis Institute of Transportation Studies (funded by State of California Air Resources Board), $21,075. Subcontract form Principal Investigator David Bunch.

4/08 - 9/09 University of California Transportation Centers (funded by the U.S. Department of Transportation and the California State Department of Transportation), $60,796.

5/07 - 3/08 University of California Transportation Centers (funded by the U.S. Department of Transportation and the California State Department of Transportation), Co-Principal Investigator with K Van Dender. $12,414.

6/06 – 9/07 California Department of Transportation, Co-Principal Investigator with T. Golob and W. Recker. $161,157.

10/05 – 9/08 U.S. National Science Foundation, Co-Principal Investigator with K. Kraemer et. al., $550,000.

8/04 – 12/05 California Department of Transportation (PATH), Co-Principal Investigator with T. Golob and W. Recker. $133,025.

8/01 - 7/02 University of California Transportation Centers (funded by the U.S. Department of Transportation and the California State Department of Transportation), Co-Principal Investigator with T. Golob. $65,057.

8/98 – 7/99 University of California Transportation Centers (funded by the U.S. Department of Transportation and the California State Department of Transportation), Co-Principal Investigator with T. Golob. $58,550.

2/96 - 11/97 PHS-NIH National Cancer Institute, Co-Principal Investigator with A.Lee-Feldstein, P. Feldstein, F. Meyskens, J. Butler, and T. Buchmueller. $237,147.

7/96 - 6/97 University of California Energy Institute, Co-Principal Investigator with D. Bunch,T. Golob. $25,000.

3/95 - 12/96 California Energy Commission and Southern California Edison Co., Co-Principal Investigator with D. Bunch,T. Golob. $997,655.

Referee : *American Economic Review, Econometric Reviews, Econometrica, The Econometrics Journal, Empirical Economics, Energy Economics, International Economic Review, Journal of Applied Econometrics, Journal of Business and Economic Statistics, Journal of Choice Modeling, Journal of Econometrics, Journal of Urban Economics, Mathematical Social Science, The Rand Journal of Economics, Regional Science and Urban Economics, Research in Transportation Economics, Review of Economics and Statistics, Scandinavian Journal of Economics, Statistical Papers, Transportation Research (A, B, B, D, E), Transport Reviews, Transportation, Transportation Science, Advances in Econometrics, Journal of Regional Science,* MacMillan Publishing Co., M.I.T. Press, National Science Foundation, North-Holland Publishing Co., Brooks/Cole Publishing Co., Pearson Publishing Company, University of Rhode Island Transportation Center, Social Science Research Council of Canada, U.C. Energy Research Groups, U.C. Transportation Centers, U.S. – Israel Binational Science Foundation, Research Grants Council of Hong Kong, Transportation Research Board.

Editorial Advisory Board Member: T*ransportation Research B, Economics of Transportation, Journal of Choice Modeling.*

Tenure reviews: Stanford University, UC Berkeley, USC, Cornell, University of Oklahoma, Dickerson College

Recent Seminar Presentations: (all Department of Economics, unless otherwise noted): University of California, Riverside (2015), International Association of Travel Behavior Modelers, Austin, Texas (2015), International Choice Modeling Conference, Windsor, England (2015), Department of Transportation, Danish Technical University (2014), ZEW, University of Mannheim (2014), U.S. Office of Controller of Currency (2013), University of California, San Diego (2013), International Choice Modeling Conference, Sydney, Australia (Plenary Speaker, 2013), University of Texas, Austin (Civil and Environmental Engineering, 2013), International Association of Travel Behavior Modelers, Toronto, Canada (2012), UC Davis Institute of Transportation Studies (2012), UC Berkeley Institute of Transportation Studies (2011), Econometric Society World Congress, Shanghai, China (2010), International Association of

Travel Behavior Modelers, Jaipur, India (2009), U. C. Santa Barbara (2008), International Conference on Design and Public Policy: Markets for Congestion and Carbon Trading. University of Essex, UK (2007), International Association of Travel Behavior Modelers, Kyoto, Japan, (2006), University of Oregon (2005), Institute of Transportation Research, University of California, Irvine (2005), University of Southern California (2004), Free University of Amsterdam (2003), University of Indiana (2002), University of Illinois, Champaign (2002), U. C. Berkeley (2001), RAND Statistics Seminar (2000), U.C. Davis (1999), Portland State University (1999), Claremont Graduate School (1998), Northwestern University (1998), U.S. Federal Reserve Board of Governors (1998), Oak Ridge National Laboratory (1998).

University and Campus Service:

UNIVERSITY WIDE:
Coordinating Committee on Graduate Affairs May 2002 – Sept. 2003.
Chair of External Advisory Committee, University of California Energy Institute, 2007-2011.

CAMPUS :
Executive Committee, UCI Institute of Transportation Studies, 2006 - .
UCI Academic Planning Group, 2015 – 2016.
UCI Committee on Academic Personnel, 2011 – 2014 (Chair 2013-2014).
UCI Academic Senate Cabinet 2002-2003, 2013-2014, and 2015 - 2016.
UCI School of Social Science Ad Hoc Committee on Graduate Student Support, Winter 2005.
Chair, Graduate Council 2002-2003 and 2015 - 2016
Vice-Chair, Graduate Council 2001-2002
Graduate Council member and housing subcommittee 2003 - 2004
UCI Ad-Hoc Committee on Graduate and Family Housing, 2002 and 2003-2004
UCI Academic Senate Representative Assembly 1998 – 2000 and 2010 - 2012.
Policy Advisory Committee, UCI Institute of Transportation Studies, 1993-2002
Executive Committee, UCI Institute for Mathematical Behavioral Sciences, 1993-2002
Academic Senate Computer Policy Committee, 1994-1996
UCI Ad-Hoc Committee on Survey Research, 1992-1993
UCI Campus Committee on Statistics, 1992-1993

ECONOMICS DEPARTMENT:
Chair, July 2005 - .June 2010.
Econometrics Recruiting Chair, 2015 - 2016
Director of Graduate Studies, 1998-2004, 1993-1996, and 1987-1990.
Graduate Committee, 1984 – 2005.
Faculty Recruiting Co-Chair, 2003-2004 and 1994-1995.
Chair, Ad Hoc Departmental Committee to Define Recruiting Focus, Summer, 2002.
Numerous Departmental Personnel committees

3. Dr. George Parsons

CURRICULUM VITAE

GEORGE PARSONS

Professor, School of Marine Science & Policy, University of Delaware
Joint Appointments in Departments of Economics and Applied Economics and Statistics

Address 204 Robinson Hall, University of Delaware, Newark, DE 19716
gparsons@udel.edu

Education A.B., Economics, Indiana University, 1975
Ph.D., Economics, University of Wisconsin, 1984

Employment Economist, US Environmental Protection Agency, 1984-1985
Assistant, Associate & Full Professor; School of Marine Science & Policy*; College of Earth, Ocean, and Environment; University of Delaware; 1986-present.
*Formerly the College of Marine Studies and the College of Marine and Earth Studies

Other Appointments

Director of Marine Policy; School of Marine Science & Policy*; College of Earth, Ocean, and Environment; University of Delaware, 2004 – 2014
Sabbatical, University of Colorado, Environment & Behavior Program, 1992
Special Assistant to Milton Russell; Assistant Administrator for Policy Planning and Evaluation; US Environmental Protection Agency; Washington, DC, 1985
*Formerly the College of Marine Studies and the College of Marine and Earth Studies

Editorial

Associate Editor, *Journal of Environmental Economics and Management*, 1994-1997
Associate Editor, *Marine Resource Economics,* 1998-2008

Other Selected Significant Professional Activities:

Board Member, Society for Benefit-Cost Analysis, 2016
Member, US EPA Science Advisory Board, Economics Subcommittee, 2008 - 2012
National Resource Council, National Academy of Sciences, Committee on Sampling and Surveying of Marine Recreational Fishing, 2005-6
Board of Directors of the Association of Environmental and Resource Economists, 1999 – 2001
Panel to Evaluate the Economics of Erosion Hazards in Coastal Areas, The H. John Heinz III Center for Science, 1998
Environmental Economics Forum: How Can Researchers and EPA Work Together to Improve Benefit-Cost Analysis? United States Environmental Protection Agency, Massachusetts Institute of Technology, Cambridge, MA, 1997
National Forum on Environment and Natural Resources R&D for the President's National Science and Technology Council, Washington, D.C., 1993

Publications (Most articles can be downloaded at http://works.bepress.com/george_parsons/)

Fooks, J., Messer, K., Duke, J., Johnson, J., Li, T. & Parsons, G. (2017) Tourist viewshed externalities and wind energy production. *Agricultural and Resource Economics Review.*

Parsons, G. (2017). Travel cost models. In Champ, P., Boyle, K., and Brown, T. (Eds.) A Primer on Non-Market Valuation. Second Edition. Dordrecht: Springer.

Myer, K., Parsons, G. & Train, K. (2017). Inadequate response to frequency of payments in contingent valuation of environmental goods. In McFadden, D. & Train, K. (Eds.) Contingent Valuation of Environmental Goods: A Comprehensive Critique. Edward Elgar Publishing Ltd.

Parsons, G. & Myers, K. (2016) Fat tails and truncated bids in contingent valuation: an application to an endangered bird species. *Ecological Economics*, 129, 210-219. {Reprinted in McFadden, D. & Train, K. (Eds.) (2017) Contingent Valuation of Environmental Goods: A Comprehensive Critique. Edward Elgar Publishing Ltd.}

Hidrue, M. and Parsons, G. (2015). Is there a near-term market for vehicle to grid electric vehicles? *Applied Energy*, 151: 67-76.

Johnston, R., Parsons, G., & Ramachandran M. (2015). Benefits transfer combining revealed and stated preference data: nourishment and retreat options for the Delaware Bay beaches. In Johnston, R., Rolfe, J., Rosenberger, R., & Brouwer, R. (ed.) (2015) Benefit Transfer of Environmental and Resource Values: A Handbook for Researchers and Practitioners. Amsterdam: Springer.

Parsons, G., Hidrue, M., Kempton, W., & Gardner, M. (2014). Willingness to pay for vehicle-to-grid (V2G) electric vehicles and their contract terms. *Energy Economics,* 42, 313-324.

Parsons, G. (2014). Nonmarket valuation. In Haab T. & Whitehead J. (eds.) Environmental and Natural Resource Economics: An Encyclopedia, *(pp. 230-2). Oxford: Greenwood.*

Parsons, G., Chen, Z, Hidrue, M., Standing, N., & J. Lilley. (2013). Valuing beach width for recreational use combining revealed and stated preference data. *Marine Resource Economics, 28(3), 221-241.*

Parsons, G. (2013). Travel cost methods. In Shogren, J. The Encyclopedia of Energy, Natural Resource, and Environmental Economics, volume 3, pp. 349-358, Amsterdam: Elsevier, England.

Hidrue, M., Parsons, G., Kempton, W., & Gardner, M. (2011). Willingness to pay for electric vehicles and their attributes. *Resource and Energy Economics,* 33, 686-705.

Krueger, A., Parsons, G., & Firestone, J. (2011). Valuing the visual disamenity of offshore wind projects at varying distances from the shore: an application to Delaware's shoreline. *Land Economics,* 87(2), 268-283.

Edwards, P., Parsons, G., & Myers, K. (2011). The economic value of viewing migratory shorebirds on the Delaware Bay: an application of the single site travel cost model using on-site data. *Human Dimensions of Wildlife,* 16, 435-444.

Huang, J., Parsons, G., Poor, J., & Zhao, M. (2011). Combined conjoint-travel cost demand model for measuring the impact of erosion and erosion control programs on beach recreation. In J. Whitehead, T. Haab, and J.-C. Huang (Eds.), *Preference Data for Environmental Valuation: Combining Revealed and Stated Preference Approaches.* New York: Routledge: 115-38.

Parsons, G. & Stefanova, S. (2011). Gauging the value of short-term site closures in a travel-cost rum model of recreation demand with a little help from stated preference data. In J. Whitehead, T. Haab, and J.-C. Huang (Eds.), *Preference Data for Environmental Valuation: Combining Revealed and Stated Preference Approaches* (pp. 239-252). New York: Routledge.

Haab, T., Whitehead, J., Parsons, G., & Price, J. (2010). Effects of information about invasive species on risk perception and seafood demand. *Resource and Energy Economics,* 32, 586-599.

Parsons, G. & Kang, A. (2010). Compensatory restoration in a random utility model of recreation demand. *Contemporary Economic Policy,* 28(4), 453-463.

Meyers, K., Parsons, G., & Edwards, P. (2010). Measuring the recreational use value of migratory shorebirds: a stated preference study of birdwatching on the Delaware Bay. *Marine Resource Economics,* 25, 247-264.

Abrams, B. & Parsons, G. (2010). The uneasy case for subsidizing energy efficiency. *The Milken Institute Review*, 12(1), 63-70.

Parsons, G., Kang, A., Leggett, C., & Boyle, K. (2009). Valuing beach closures on the Padre Island National Seashore. *Marine Resource Economics*, 24, 213-235.

Abrams, B. & Parsons, G. (2009). Is CARS a clunker? *The Economist's Voice,* 2009.

Parsons, G. & Thur, S. (2008). Valuing the changes in the quality of coral reef ecosystems in the Caribbean: a stated preference study of recreational SCUBA diving in the Bonaire National Marine Park. *Environmental and Resource Economics*, 40, 593-608.

Borchers, A., Duke, J., & Parsons, G. (2007). Does willingness to pay for green energy differ by source? *Energy Policy,* 35(6), 3327-3334.

Parsons, G., Morgan, A., Whitehead, J., & Haab T. (2006). The welfare effects of pfiesteria-related fish kills: a contingent behavior analysis of seafood consumers. *Agricultural and Resource Economics Review*, 35(2), 349-356.

Reiter, M., Scarborough, B., Fan, C., Parsons, G., & Thur, S. (2006). An interdisciplinary model for the St. Jones River Watershed, Delaware: development, results, and implications. *Journal of Environmental Monitoring and Restoration,* 2, 38-50.

Parsons, G. & Noailly, J. (2004). A value capture tax for financing beach nourishment projects: an application to ocean beaches in Delaware. *Ocean and Coastal Management,* 47, 49-61.

Scrogin, D., Boyle, K., Parsons, G., & Plantinga, A. (2004). Effects of regulations on expected catch, expected harvest, and site choice of recreational anglers. *American Journal of Agricultural Economics*, 86(4), 963-974.

von Haefen, R., Pheneuf, D., & Parsons, G. (2004). Estimation and welfare analysis with large demand systems. *Journal of Business and Economics Statistics,* 22(2), 194-205. {Reprinted in Herriges, J. & Kling, C. (Eds.) *Revealed Preference Approaches to Environmental Valuation: Volume II,* Aldershot: Ashgate Publishing, 2008.}

Whitehead, J., Haab, T., & Parsons, G. (2003). Economic effects of pfiesteria. *Ocean and Coastal Management*, 46, 845-858.

Parsons, G. (2003). The travel cost model. In Champ, P., Boyle, K., and Brown, T. (Eds.) *A Primer on Non-market Valuation* (pp. 269-329), London: Kluwer Academic Publishing.

Parsons, G. & Massey, D. (2003). A random utility model of beach recreation. In Hanley, N., Shaw, W., & Wright, R. (Eds.). *The New Economics of Outdoor Recreation* (241-267), Edward Elgar. 2003.

Wakefield, J. & Parsons, G. (2003). A comparison of nourishment and retreat costs on Delaware's ocean beaches. *Shore and Beach*, 71(1), 23-26.

Parsons, G. (2001). The random utility model for valuing recreation uses of the environment. In *Societal Values in Public Policy*, OECD.

Parsons, G. & Powell, M. (2001). Measuring the cost of beach retreat. *Coastal Management,* 29, 91-103.

Hauber, B. & Parsons, G. (2000). The effect of nesting structure specification on welfare estimation in random utility model of recreation demand: an application to the demand for recreational fishing. *American Journal of Agricultural Economics,* 82(3), 501-514.

Parsons, G., Plantinga, A., & Boyle, K. (2000). Narrow choice sets in random utility models of recreation demand.

Land Economics, 76(1), 86-99.

Parsons, G., Massey, D., & Tomasi, T. (1999). Familiar and favorite sites in a random utility model of beach recreation. *Marine Resource Economics*, 14, 299-315. {Reprinted in Herriges, J. & Kling, C. (Eds.) *Revealed Preference Approaches to Environmental Valuation: Volume II,* Aldershot: Ashgate Publishing, 2008.}

Parsons, G., Jakus, P., & Tomasi, T. (1999). A Comparison of welfare estimates from four models for linking seasonal recreational trips to multinomial models of site choice. *Journal of Environmental Economics and Management,* 38(2),143-157.

Parsons, G. & Hauber, A. (1998). Spatial boundaries and choice set definition in a random utility model of recreation demand. *Land Economics,* 74(1), 32-48.

Parsons, G. & Wilson, A. (1997). Incidental consumption in recreation demand. *Agricultural and Resource Economics Review,* 26(1): 1-6.

Parsons, G. & Kealy, M. (1995). A demand theory for number of trips in a random utility model of recreation. *Journal of Environmental Economics and Management,* 29(3), 357-367.

Parsons, G. & Kealy, M. (1994). Benefits transfer in a random utility model of recreation. *Water Resources Research,* 30(8), 2477-2483.

Parsons, G. & Needelman, M. (1992). Site aggregation in a random utility model of recreation. *Land Economics* , 68(4), 418-33. . {Reprinted in Herriges, J. & Kling, C. (Eds.) *Revealed Preference Approaches to Environmental Valuation: Volume II,* Aldershot: Ashgate Publishing, 2008.}

Parsons, G. & Kealy, M. (1992). Randomly drawn opportunity sets in a random utility model of lake recreation. *Land Economics,* 68(1), 93-106.

Desvouges, W., Naughton, M., & Parsons, G. (1992). Benefits transfer: conceptual problems in estimating water quality benefits using existing studies. *Water Resources Research,* 28(3), 675-683.

Parsons, G. (1991). The effect of coastal land use restrictions on housing prices: a repeat sale analysis. *Journal of Environmental Economics and Management,* 22(1), 25-37.

Parsons, G. & Wu, Y. (1991). The opportunity cost of coastal land-use controls: an empirical analysis. *Land Economics,* 67(3), 308-316.

Parsons, G. (1991). A note on choice of residential location in travel cost demand models. *Land Economics,* 67(3), 360-364.

Parsons, G. (1990). Hedonic prices and public goods: an argument for weighting locational attributes in hedonic regressions by lot size. *Journal of Urban Economics,* 27(3), 308-321.

Parsons, G. (1989). On the value of the condition of a forest stock: comment. *Land Economics,* 65(1), 68-72.

Parsons, G. (1987). The opportunity costs of residential displacement due to coastal land use restrictions: a conceptual framework. *Marine Resource Economics,* 4(2), 111-122.

Parsons, G. (1986). An almost ideal demand system for housing attributes. *Southern Economic Journal,* 53(2), 347-363.

Milon, W. & Parsons, G. (1986). Measuring the site specific recreation benefits resulting from improved water quality: an upper bound approach. *Water Resources Bulletin,* 22(1), 133-134.

Book Reviews

Parsons, G. (2001). Review of Herriges, J. & Kling, C. (Eds.) *Valuing Recreation and the Environment*. Cheltenham, UK: Edward Edgar. In *American Journal of Agricultural Economics*, 83 (4), 1101-1105.

Parsons, G. (1998). Short review of Morgenstern, R. (Ed.) (1998). *Economic Analyses at EPA*, Washington DC: Resources for the Future. In *Environment.*

Parsons, G. (1995). Review of Bromley, D. (Ed.) (1996). *Making the Commons Work: Theory, Practice and Policy*. In *Southern Economic Journal*, 61(4), 1253-1255.

Parsons, G. (1994). Review of Penning-Rowsell, E., Green, C., Thompson, P., Coker, A., Tunstall, S., Richards, C. & Paker, D. (1992). *The Economics of Coastal Management: A Manual of Benefit Assessment Techniques*. London: Belhaven Press. In *Ocean and Coastal Management* 24(3), 215-24.

Research Grants

National Science Foundation, "Toward Sustainable Urban Estuaries in the Anthropocene" $752,577, 2013-2018 (co-PI with C. Sommerfield).

Bureau of Ocean Energy Management, "Atlantic Offshore Wind Energy Development: Public Attitudes, Values, and Implications for Tourism and Recreation," $200,000, 2012-2014, (Co-PI with J. Firestone).

National Oceanic and Atmospheric Administration, Sea Grant, "The Economic Impact of Offshore Wind Farms on Beach Use in the Mid-Atlantic Region," $140,000, 2012-2014.

Department of Energy, "Electric vehicles for vehicle-to-grid power," Sub-Investigator on economics and marketing of electric vehicles, $175,000, 2008-2010 (my share).

National Oceanic and Atmospheric Administration, Sea Grant, "An Economic Analysis of Shorebirds and Shorebirding in the Mid-Atlantic," $155,000, 2009-2011. (continuation of 2007-2008 grant).

National Oceanic and Atmospheric Administration, Sea Grant, "An Economic Analysis of Shorebirds and Shorebirding in the Mid-Atlantic," $160,000, 2007-2008.

US Environmental Protection Agency, "Calibration in Benefits Transfer" Sub-Investigator, $25,000, 2006-8 (my share with no-cost extensions).

US Environmental Protection Agency, "Bayesian and Function Transfers in Benefit Assessment," Sub-Investigator, $30,000, 2006-8 (with no-cost extensions).

National Oceanic and Atmospheric Administration, Sea Grant, "An Economic Analysis of Shorebirds and Shorebirding in the Mid-Atlantic," $160,000, 2007-2008.

Coastal Response Research Center, "Monetary Values and Restoration Equivalents for Lost Recreational Services on the Gulf Coast of Texas Due to Oil Spills and Other Environmental Disruptions," $133,326, 2006-9.

National Oceanic and Atmospheric Administration, Sea Grant, "An Economic Analysis of Mid-Atlantic Beach Use," $145,528, 2005-6.

Inter-American Development Bank, "Review of Environmental Valuation Studies for Coastal Resources in Latin America and the Caribbean" (co-PI with Lee G. Anderson), $130,000, 2003

National Oceanic and Atmospheric Administration, Environmental Cooperative Science Center, "Ecological Model of the St. Jones Watershed", Through Michael Reiter, Delaware State University, $40,000, 2001-2004

National Park Service, "Baseline Valuation Studies of Park Visitor Activities" Co-PI with Industrial Economics Inc.,

(co-PI with Kevin Boyle, John Duffield, and Nancy Bockstael), $500,000, 1999-2001

National Oceanic and Atmospheric Administration, "The Economic Effects of Pfiesteria in the Mid-Atlantic Region," (co-PI with Tim Haab, John Whitehead, Doug Lipton, and Jim Kirkley), $50,000, 1999-2001

U.S. Environmental Protection Agency, "The Benefits of Water Quality Improvements Under the Clean Water Act," Through a Research Triangle Institute Cooperative Agreement, $25,000, 1998-2000
National Oceanic and Atmospheric Administration, Sea Grant, "Nourishment versus Retreat: An Economic Analysis of Beach Management in Delaware," (co-PI with Tony Dalrymple). $100,000, 1999-2001

National Oceanic and Atmospheric Administration, Sea Grant, "The Recreation Value of Coastal Beaches," (co-PI with Ted Tomasi), $50,000, 1995-97

U.S. Environmental Protection Agency, Cooperative Agreement, "Methods Development for Measuring the Economic Benefits of Environmental Improvements," $275,000, 1992-1997

U.S. Environmental Protection Agency, Cooperative Agreement, "Measuring the Benefit of Water Quality Improvements," $155,000, 1988-1992

National Oceanic and Atmospheric Administration, Sea Grant, "The Economics of Beach Preservation," $74,000, 1989--1993

National Oceanic and Atmospheric Administration, Sea Grant, "An Applied Welfare Analysis of Coastal Land Use Restrictions," $79,000, 1987-1989

Presentations

Invited talks: University of Rhode Island, University of Colorado, University of Washington, North Carolina State University, Appalachian State University, Resources for the Future, US Environmental Protection Agency, National Oceanic and Atmospheric Administration, University of Maryland, Industrial Economics Inc., University of Central Florida, Franklin and Marshall, Charles Sturt (Australia), and University of Maine.

Conferences I frequent: Allied Social Science Association, Association of Environmental and Resource Economists, Southern Economics Association, Northeastern Economics Association, W-133 Regional Meetings, Western Agricultural Economics Association, Resource for the Future Workshop, Agricultural and Applied Economics Association, Law Seminars International, Oil Spill and Ecosystem Science, and Society for Benefit-Cost Analysis.

Review & Referee

Journals: Agricultural and Resource Economics Review, American Economic Review, American Journal of Agricultural Economics, Cahiers d'Economie et Sociologie Rurales, Climate Change, Coastal Management, Contemporary Economic Policy, Eastern Economics Journal, Economic Inquiry, Ecological Economics, Environmental Management, Environmental & Resource Economics, Environment & Planning, Journal of Agricultural Economics, Journal of Applied of Econometrics, Journal of Econometrics, Journal of Environmental Economics & Management, Journal of Environmental Management, Journal of Human Resources, Journal of Public Economics, Land Economics, Leisure Science, Marine Resource Economics, Ocean & Coastal Management, Oxford Economic Papers, Resource and Energy Economics, Review of Economics and Statistics, Southern Economic Journal, Transportation Research Board, and Water Resources Research.

Proposal review panels: US Environmental Protection Agency, National Oceanic and Atmospheric Administration, Department of Natural Resources and Environmental Control, DE, the National Science Foundation, and the Israel Science Foundation.

Paper selection committee Association of Environmental & Resource Economists, American Agricultural Economics Association, Northeastern Agricultural Economics Association, and, Society for Benefit-Cost Analysis

Consultancies

Abt Associates, Arnold & Porter, Delaware Department of Natural Resources & Environmental Control, Delaware Bay Program, Cardno-Entrix Inc., Industrial Economics, New York Department of State, Massachusetts Bay Program, Research Triangle Institute, Stratus Consulting, Triangle Economic Research, and US Environmental Protection Agency.

Professional Societies (not all current)

American Economic Association, Association of Environmental & Resource Economists
American Agricultural Economic Association, Northeastern Agricultural & Resource Economics Association, and Society for Benefit-Cost Analysis.

Courses Taught (all graduate level)

Topics in Environmental Economics, Valuing the Environment, Environmental Economics, Applied Policy Analysis, Marine Policy Seminar, Introduction to Marine Policy, and Benefit-Cost Analysis.

Student Advisement

Major Advisor to Students Receiving a PhD (16 Completed)
Lingxao (Alfred) Yan – PhD Marine Studies (current)
Kate Karpov - PhD Economics (current)
Yossi Shirazi - PhD Marine Studies (current)
Ed Carr - PhD Marine Studies (current)
Zoe Chen - PhD Economics, 2016
Jenna Toussaint - PhD Economics, 2015
Kelley Myers - PhD Marine Studies 2013 – ***E. Sam Fitz Award in SMSP 2013***
Michael Hidrue - PhD Economics 2011 - ***Best PhD Dissertation in Economics 2011***
Georgi Spiridonov - PhD Economics 2010
Stela Stephanova - PhD Economics 2009 - ***Best PhD Dissertation in Economics 2010***
Peter Edwards - PhD Marine Studies 2009
Ami Kang,PhD - PhD Marine Studies 2009
Christy Loper - PhD Marine Studies 2008
Nasser Yayi - PhD Economics 2007
Steve Thur - PhD Marine Studies 2003
Eric Helm - PhD Economics 2003
Ashton Morgan - PhD Economics 2002
D. Mathew Massey - PhD Economics 2002
Jeffery R. Wakefield - PhD Economics 2001
A. Brett Hauber IV - PhD Economics 2000

Major Advisor to Students Receiving a Masters (34 Completed)
Meredith Kurz, MMP Marine Studies (current)
Jeff Flood, MPP Marine Studies 2016
Yangpeiun Xu, MPP Marine Studies 2016
Harrison Celone - MA Economics 2012
Mariel Montoney - MA Economics 2012

Naomi Standing - MS Economics 2011
Semere Meshesha - MA Economics 2011
Royce Vaughn - MA Economics 2011
Laura Ratz - MA Economics 2010
Christina Moore - MA Economics 2010
Kelley Appleman - MS Economics 2009 - ***Best MS Thesis in Economics 2010***
Laura Magee - MS Economics 2008
Amanda Wenzel - MMP Marine Studies 2008
Andrea Geiger - MMP Marine Studies 2006
Jeffery Levinson - MMP Marine Studies 2003
Heather Daniel - MMP Marine Studies 2000
Ashton Morgan - MA Economics 2000
Joelle Noailly - MS Economics 1999
Jeffery Wakefield - MA Economics 1999
Erik Helm - MA Economics 1999
Kelly L. Bungee-Rogers - MMP Marine Studies 1998
Marianna Mittelstadt - MA Economics 1998
Pravitha Beemsen - MA Economics 1997
Anupit Supnithadnaporn - MMP Marine Studies 1996
Wayne Gibson - MA Economics 1996
Susan Flaherty - MA Economics 1996
Mariead Kenny - MA Economics 1996
Aaron Wilson - MA Economics 1996
Michael Powell - MMP Marine Studies 1995
Robert H. Boyles, Jr. - MMP Marine Studies 1993
Sean Carmen - MA Economics 1994
Debbie Birnbaum - MA Economics 1993
Michael Needelman - MS Economics 1991
Terry Parker - MMP Marine Studies 1990
Becky Metzner - MMP Marine Studies 1988

Major Advisor to Students Receiving an Honors BS:
Leigh Robinson – BS Economics, 2016
Julia Eve Kratzer - BS Economics 1994

Outside Reader for students earning degrees at other Universities:
Weibin Xu, PhD Economics, Virginia Tech, 2016
Tom Andrews, PhD Economics, Temple University, 1996
Martin van Bueren, PhD Economics, The University of Western Australia, 1999

Committee Member or Reader on over 40 other PhD and Master Degree Students

College and University Service
Promotion and Tenure Committee, 2016-8
Chair, Search Committee for Assistant Professor of Marine Policy, 2016-7
Member, Search Committee for Assistant Professor of Applied Economics, 2014-5
Chair, Search Committee for 4 New School-Wide Professors, 2014 -5
Program Director, Marine Policy, College of Earth, Ocean, and Environment, 2004 - 2014
Chair, Search Committee for Assistant Professor of Marine Policy, 2012-13
Chair, Promotion and Tenure Committee, College of Earth, Ocean, and Environment, 2012-14
Search Committee for Cosgrove Professor of the Environment, 2010-13
College Transition Committee, College of Marine and Earth Studies, 2009
Committee on Coastal Community Development, University of Delaware, 2005
Search Committee for Director of Business Office, College of Marine Studies, 2005
Chair, Promotion and Tenure Committee, College of Marine Studies, 2003-4

Advisory Council on International Programs, University of Delaware, 2001-2004
Committee on the Formation of a One-Year Degree Program, 2003
Acting Chair, Marine Policy Program, Graduate College of Marine Studies, 2002
Committee on Research, University of Delaware, 1999 – 2001
Search Committee for Assistant Professor of Marine Policy, 2000-2001
Chair, Search Committee for Assistant Professor of Marine Policy, 1999-2000
Task Force on Faculty for the Middle-States 10-year Accreditation Process, University of Delaware, 1999 – 2000
Chair, Search Committee for Chief Executive Officer, College of Marine Studies, 1998-1999
Graduate Committee, Department of Economics, 1988 – 1990, 1994
Faculty Senator, University of Delaware, 1989-1991
Search Committee for Program Leader for the Marine Advisory Service, College of Marine Studies, 1991
Computer Committee, Graduate College of Marine Studies, 1987 – 1990
Awards Committee, Graduate College of Marine Studies, 1989

Appendix B. Conflicts of Interest Declarations

1. Dr. David Bunch

ORGANIZATIONAL CONFLICT OF INTEREST CERTIFICATE

Customer: U.S. Environmental Protection Agency

Contractor: ICF Incorporated, LLC, 9300 Lee Highway, Fairfax, VA 22031

Prime Contract: EP-C-16-020, Work Assignment 0-16

Subject Report: Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge

Subcontractor/Peer Reviewer: David Bunch, University of California at Davis

In accordance with EPAAR 1552.209-70 through 1552.209-73, Subcontractor/Consultant certifies to the best of its knowledge and belief, that:

✓ No actual or potential conflict of interest exists.

_____ An actual or potential conflict of interest exists. See attached full disclosure.

Subcontractor/Consultant certifies that its personnel, who perform work on this contract, have been informed of their obligations to report personal and organizational conflict of interest to Contractor and Subcontractor/Consultant recognizes its continuing obligation to identify and report any actual or potential organizational conflicts of interest arising during performance under referenced contract.

David R Bunch

Subcontractor/Consultant

Sept. 8 2017

Date

100 Cambridgepark Drive, Suite 500 ····· Cambridge MA 02140 ····· 617.250.4200 ····· 617.250.4281 fax ····· www.icfi.com

2. Dr. David Brownstone

ORGANIZATIONAL CONFLICT OF INTEREST CERTIFICATE

Customer:	U.S. Environmental Protection Agency
Contractor:	ICF Incorporated, LLC, 9300 Lee Highway, Fairfax, VA 22031
Prime Contract:	EP-C-16-020, Work Assignment 0-16
Subject Report:	Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge
Subcontractor/Peer Reviewer:	David Brownstone, University of California at Irvine

In accordance with EPAAR 1552.209-70 through 1552.209-73, Subcontractor/Consultant certifies to the best of its knowledge and belief, that:

____X___ No actual or potential conflict of interest exists.

_____ An actual or potential conflict of interest exists. See attached full disclosure.

Subcontractor/Consultant certifies that its personnel, who perform work on this contract, have been informed of their obligations to report personal and organizational conflict of interest to Contractor and Subcontractor/Consultant recognizes its continuing obligation to identify and report any actual or potential organizational conflicts of interest arising during performance under referenced contract.

_____
Subcontractor/Consultant

_October 6, 2017____
Date

100 Cambridgepark Drive, Suite 550 ..... Cambridge MA 02140 ..... 617.250.4200 ..... 617.250.4201 fax ..... www.icf.com

3. Dr. George Parsons

ORGANIZATIONAL CONFLICT OF INTEREST CERTIFICATE

Customer:	U.S. Environmental Protection Agency
Contractor:	ICF Incorporated, LLC, 9300 Lee Highway, Fairfax, VA 22031
Prime Contract:	EP-C-16-020, Work Assignment 0-16
Subject Report:	Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge
Subcontractor/Peer Reviewer:	George Parsons, University of Delaware

In accordance with EPAAR 1552.209-70 through 1552.209-73, Subcontractor/Consultant certifies to the best of its knowledge and belief, that:

X No actual or potential conflict of interest exists.

_____ An actual or potential conflict of interest exists. See attached full disclosure.

Subcontractor/Consultant certifies that its personnel, who perform work on this contract, have been informed of their obligations to report personal and organizational conflict of interest to Contractor and Subcontractor/Consultant recognizes its continuing obligation to identify and report any actual or potential organizational conflicts of interest arising during performance under referenced contract.

George Parsons
Subcontractor/Consultant

16-1-2017
Date

100 Cambridgepark Drive, Suite 500 — Cambridge MA 02140 — 617.250.4200 — 617.250.4261 fax — www.icfi.com

Appendix C. Peer Reviewer Mid-review Meeting Notes

SUMMARY NOTES:
PEER REVIEW FOR CONSUMER WILLINGNESS TO PAY FOR VEHICLE ATTRIBUTES
MID-REVIEW MEETING
September 21, 2017 – 4:00 PM

Attendees:

Gloria Helfand, EPA
David Brownstone, University of California at Irvine
David Bunch, University of California at Davis
George Parsons, University of Delaware
Jon Hecht, ICF
Andie Fritz, ICF

Welcome, Introductions, and Roles

Jon Hecht gave an overview of the peer review program generally, and as applies to the report, "Consumer Willingness to Pay for Vehicle Attributes: What is the Current State of Knowledge?"

All participants then gave brief introductions of themselves and their backgrounds.

Jon Hecht then confirmed that the reviewers received their technical packages. Each package included a charge letter, conflict of interest form, and report for review.

Overview of the Project and Report

Gloria Helfand gave a brief overview of the project and report. EPA has been developing Willingness to Pay (WTP) estimates for a variety of vehicle characteristics. Characteristics of interest included fuel economy, performance, acceleration, etc. Estimates were derived from demand models of existing published light-duty vehicle demand studies. If vehicle standards that EPA issues affect these other vehicle characteristics, having estimated values for these attributes might enable EPA to develop better estimates of the benefits and costs of its standards.

Open Discussion and Questions

The meeting then turned to open discussion to answer the reviewers' questions.

David Bunch asked how EPA plans to use the WTP estimates in the future. He also mentioned changes in fuel economy, performance, and retail vehicle costs. Sales shifts and their implications should be considered. He commented that the WTP results are quite broad and asked how he should focus his attention when providing comments. Gloria Helfand agreed that sales shifts should be considered, but there is no consensus on the effects of vehicle sales. The studies show a wide variation in WTP calculations for fuel economy, performance, etc. Gloria

Helfand encouraged the reviewers to provide suggestions for meta-analysis. She also mentioned a follow-on project to analyze these wide variations in numbers and identify the differences. It is important that the existing studies and WTP estimates identified in this study be reasonable as they will be considered in this following study.
David Bunch also suggested direct tradeoff comparisons of performance and fuel economy. This may be more useful than considering WTP, as the price coefficient can be difficult to determine. George Parsons also suggested another simulation. He suggested having two sets of demand model vehicles, one set before the standards and one set after the standards. Certain attributes should then be changed to determine how consumer welfare and choice are affected.

David Brownstone reiterated that the estimates vary greatly. He suggested obtaining more data and establishing a confidence around WTP. Gloria Helfand responded that the study has considered a variety of existing literature. The study attempted to see if there is any consensus forming across them. The reviewers should comment if they feel that a certain study is significantly more accurate or if a certain study is unreasonable and should not be included.

David Bunch asked for clarification on Charge Question 1. Gloria Helfand responded that the reviewers should comment if certain studies are better than others, or if certain studies should be dropped from consideration. David Bunch also mentioned a European study that may be valuable to consider. Gloria Helfand responded that EPA and RTI decided to focus on US studies post 1995. However, if there were studies that met the criteria and were omitted, EPA welcomes these study suggestions.

David Brownstone asked about the review submission format. Andie Fritz responded that submitted comments should maintain the tabular format.

David Bunch asked if the reviewers are allowed to discuss with each other, or to distribute the focus areas amongst the group. Gloria Helfand responded that EPA is hoping to obtain independent opinions. They are not expecting a consensus amongst the reviewers and want to avoid biased comments.

David Bunch asked how $2.50 was chosen as the price of gas standard, regardless of when and where the study occurred. Gloria Helfand responded that WTP should depend on the price of gas and they attempted to make it comparable across the studies by using a consistent price.

Schedule

Jon Hecht gave an overview of the project's schedule. The reviewers provided their availability. ICF and EPA spoke offline and finalized the review deadline for Friday, October 6, 2017. All reviews will be returned to ICF (to Andie Fritz). ICF will then compile all comments and share with EPA.

Next steps

Jon Hecht asked reviewers to respond to all charge questions and maintain the tabular format. Reviewers should also submit a cover letter that includes their name, name and address of their organization, and a statement of any real or perceived conflict(s) of interest. A completed conflict of interest form is also required. All were provided in the reviewer packages. Reviewers must also provide a current CV.

During the review period, the reviewers will send any questions to ICF. ICF will forward the questions to EPA. ICF will then share all questions and responses with the entire review team.

The review deadline is **Friday, October 6, 2017**. All reviews will be returned to ICF. If any additional review time is required, the reviewer should reach out to ICF so that an extension can be coordinated. EPA wants to ensure that the report is thoroughly reviewed.