CONTINGENT VALUATION AND REAL ESTATE DAMAGE ESTIMATION

Clifford A. Lipscomb  
*Greenfield Advisors, LLC*

Max Kummerow  
*Greenfield Advisors, LLC*

Will Spiess  
*Greenfield Advisors, LLC*

Sarah Kilpatrick  
*Greenfield Advisors, LLC*

John A. Kilpatrick  
*Greenfield Advisors, LLC*

Abstract

Real estate appraisal practice stresses three traditional approaches (cost, sales comparison, and income) to value contaminated properties. This paper discusses a fourth approach: the contingent valuation method (CVM). This method is useful in the estimation of losses for contaminated properties when market data are not available or market failure is suspected. Transactions of impaired properties often represent a small sample of all properties under consideration; efficient market assumptions, such as an adequate level of information, are usually not met. In such circumstances, stated values derived from surveys may offer complementary and valid evidence of market value. Where transactional evidence and survey results differ, further analysis of market conditions may be helpful to determine which represents the most credible value conclusion.

Suppose that a buyer who recently paid $500,000 for a home puts that same home up for sale after contamination is discovered on the property. It is often the case with impaired property that no buyers can be found; or if buyers bid for the property, there is a lack of information about the specifics of the property. In such situations, value evidence from other arms-length transactions may be scarce, making it difficult for potential buyers, the seller, and real estate appraisers to determine the market value of the home. Eventually, after some period of time, assume a buyer and the seller agree on a purchase price of $200,000 for the property. In the interim, however, suppose that a real estate researcher uses the contingent valuation method (CVM) to establish an estimate of $350,000 for the property based on a description of the property and its contamination issues. That is, survey respondents express a willingness-to-pay (WTP) of $350,000 after the scenario is described in detail.

Which valuation is more reflective of the property’s value—the first $500,000 price (an actual payment based on incomplete information) or the $350,000 price obtained via the CVM (showing a stated preference based on information presented to the
survey respondent) or the second sale at $200,000 (possibly a panic sale in a market far from equilibrium)? Market value depends on information, time, and market conditions, all of which are important issues to address when dealing with contaminated properties.

Stated preference methods were initially developed for the valuation of public goods in the fields of resource and environmental economics and in market research to estimate demand and prices. Stated preference methods consist of a family of survey techniques where individual respondents answer questions about what they would be WTP or willing to accept (WTA). For example, respondents might express WTP to avoid some environmental degradation or to accept compensation for some environmental degradation that has already occurred. These methods require the use of surveys to collect data.

Mundy and McLean (1998) provide one of the first introductions to the CVM in the real estate appraisal literature. They outline the major issues associated with the method and, among other things, discuss the kinds of survey questions one might ask depending on the particular contamination situation. Despite the limitations of the CVM, even Roddewig and Frey (2006, p. 280), who are critics of the use of CVM to value real estate, admit that “There may indeed be situations in which application of contingent valuation techniques may be useful, for example, when dealing with noneconomic goods or in real estate situations involving special-purpose or limited-market properties for which there are few, if any, actual sales transactions that can be analyzed.” We contend that the potential uses of the CVM are more expansive than this. We argue that the CVM is especially applicable in situations where properties suffer from contamination or other impairment where market failure may occur for various reasons, including the lack of information and knowledge. CV surveys are particularly useful in confirming the causes of property value diminution when contamination is present (Simons and Winson-Geideman, 2005).

Market prices often can be biased due to the violation of efficient market assumptions. Behavioral economics experiments and market research reveal biases and limitations in human perceptions, information processing, and decision making. “Bubbles” and “busts” are often observed—usually in hindsight—in patterns of transactions prices. The most recent real estate crisis provides excellent insight on this issue. Irrational exuberance pushed real estate prices up to unsustainable levels. More efficient markets are now pushing prices back down to sustainable levels. In some circumstances, properly designed surveys can offer value estimates that better reflect symmetric information in the marketplace and hence values that are more in accord with definitions of market value than prices revealed in market transactions. The CVM is, therefore, an appropriate tool to help fill in the information gaps found in markets with contaminated property.

Where transaction prices differ significantly from values obtained via the CVM, the assumption should not automatically be made that the transaction price is “right” and the survey response “wrong.” Carson and Hanemann (2005, p. 824) suggest that “Rather than seeing an inherent conflict between revealed and stated preference techniques, it is more productive to view the two approaches as complementary but
having different strengths and weaknesses. Indeed it is sometimes possible and useful to combine the two approaches.” Earnhart (2001, 2002) shows how the revealed preference and stated preference approaches may be combined. Also, the International Association of Assessing Officers (IAAO) states that “alternative methods should be investigated when absence of reliable information lessens the applicability of more conventional techniques.”

The following section of the paper defines the CVM method. There is then a brief overview of the CVM literature, followed by a discussion of the potential biases associated with the method and techniques for mitigating biases and increasing the validity of WTP estimates. Next, there is a discussion of where CVM valuations and transactions prices disagree. Continuing, the Uniform Standards of Professional Appraisal Practice (USPAP) are addressed and what it implies for the CVM. The paper closes with concluding remarks.

**WHAT IS THE CONTINGENT VALUATION METHOD**

According to the World Bank Institute (2002, p. 2), the CVM is a stated preference method used to estimate “the value that a person places on a good, usually one that is not sold in markets, such as environmental quality or good health. The approach asks people directly what they are willing to pay for the good, or what they are willing to accept to give it up, rather than inferring this from observed behavior.” Researchers strive to create real-world conditions when designing a CV survey, particularly to preserve the take-it-or-leave-it (TILI) condition, which provides an incentive for respondents to state their true (not inflated or deflated) WTP or WTA for a good.

The CVM originated in studies of the economics of public goods, or goods that have value but are not traded in formal markets. National defense, for example, has no marketplace where one can purchase it. National defense is a non-excludable and non-rivalrous public good—everyone has access to it and no single person pays for it. Economists have long been concerned with how to measure economic values in decisions that affect the provision of public goods. For example, do the benefits of cleaner air (improved visibility, fewer health-related incidents, etc.) justify the costs incurred to clean up dirty air? Because markets are unable to measure the value of non-traded public goods, economists cannot rely on market prices as they normally do to determine the value of private goods. Stated preference methods are the only methods that can measure non-use values such as “existence values.”

While generally accepted in environmental and resource economics, the CVM has been a more controversial method in real estate appraisal circles. Critics argue that it is a flawed method because it has inherent biases; these same critics maintain that there is a lack of empirical evidence to support the value estimations from CV surveys. However, the broader world of economics and marketing has accepted the CVM as useful if the inherent biases of the method are mitigated through good survey design, if it is applied in appropriate situations, and if survey responses are carefully interpreted. In the valuation of public goods and benefit-cost analysis, the CVM is as well established and accepted as the traditional three approaches (cost, sales comparison, and income) to valuation are accepted in real estate appraisal practice.
When considering contaminated properties, analysts must consider the fact that these are likely places suffering from market failure. Pollution externalities have fallen upon landowners as a result of the activities of others, often with major effects on property values. Parties to transactions may be uninformed, misinformed, experiencing irrational fear that exaggerates minor risks, or experiencing irrational denial of genuine risks.

Typically, many households state that they would not buy or rent structures where environmental contamination is present at any price; but, survey evidence tells us that sometimes households continue to live in contaminated areas for a variety of reasons (e.g., “This is ‘home’;” “I can’t afford to move;” “I don’t want to find a new job in this economy in a new location”). As a result, the volume of transactions sometimes decreases steeply in areas when contamination becomes public knowledge. Under these circumstances, the established valuation methods may be compromised by lack of sufficient transactional data meeting the standard definition of “market value.” In other circumstances, the volume of transactions may not change that much, raising the concern that households new to an area may not have an awareness of the local contamination situation.

The CVM typically seeks respondents’ estimates of WTP contingent upon the details of a constructed hypothetical market transaction. A well-constructed survey related to an issue with which residents have sufficient experience can provide a realistic, balanced, and complete picture. This kind of situation enables respondents to make rational, informed, and accurate responses used to determine market value.

**Literature Review**

**Economics Literature**

A search of the EconLit database on September 9, 2010, for the term “contingent valuation,” returned 1,698 articles. Carson (2011) provides a bibliography with thousands of CVM papers spanning the last fifty years. Much methodological research has focused on ensuring that a respondent’s statement of WTP/WTA is accurate.

Several economists argue that the validity of nonmarket valuation methods (especially the CVM) has not been demonstrated sufficiently (Mitchell and Carson, 1989; Arrow et al., 1993; Portney, 1994). However, other economists have taken steps to validate the CVM by statistically testing the difference between WTP for a good as reported from a CVM survey and WTP for a good as reported from an actual or simulated market transaction. This seems reasonable for private goods like food or parking permits, but may be problematic for public goods where actual WTP may undervalue compensating variation due to the free-rider problem (World Bank Institute, 2002).

The academic literature suggests that there may be a difference in the validity of value estimates based on whether one is valuing a public good or a private good and the respondent’s degree of familiarity with the good being valued. Generally, research suggests that stated WTP responses for public goods tend to overstate the actual WTP of consumers (Kealy, Dovidio, and Rockel, 1987; Arrow et al., 1993). Carson and
Hanemann (2005) argue that economic theory should suggest to the researcher whether or not survey-based WTP estimates will diverge from actual WTP. For example, they argue that “The actual contributions should grossly underestimate actual willingness to pay because of the free-rider problem that has long been at the core of issues surrounding the provision of public goods,” (p. 915).

For quasi-public goods, like hunting permits, Bishop and Heberlein (1979) and Mitchell and Carson (1989) find no statistically significant difference between the hypothetical and actual WTP. In an early but important behavioral economics result, Bishop and Heberlein (1979) found biases in WTP versus WTA in market transactions (deer hunting permit purchases and sales), a result that contradicts the efficient markets hypothesis that buyers and sellers act rationally. More recent behavioral economics literature demonstrates that revealed preferences suffer their own set of biases.

More recent efforts have concentrated on CVM/actual market comparisons for less familiar public goods with large non-use components (Poe, Clark, Rondeau, and Schulze, 2002). For example, Seip and Strand (1992) compared hypothetical and actual environmental organization sign-ups; Brown, Champ, Bishop, and McCollum (1996) and Champ, Bishop, Brown, and McCollum (1997) compared real and hypothetical donations for road removal on the north rim of the Grand Canyon; Champ and Bishop (2001) compared hypothetical and real contributions for a wind energy project; and Poe, Clark, Rondeau, and Schulze (2002) compared hypothetical and real sign-up rates and contributions for a green electricity program in the Buffalo, NY region. These studies generally have concluded that there are considerable differences between hypothetical and actual contributions, largely attributable to the inherent biases of the CVM. One interesting point made by Poe, Clark, Rondeau, and Schulze (2002, p. 119) is that when they considered the median bid to actual value ratio, there is a possibility that “the mechanism underestimates true values by perhaps as much as 35 percent, a figure that would likely bring the real sign-up rate much closer to the sign-up rates...obtained from the [discrete choice] responses.” In summarizing the impact of these studies on methodological practices, they admit that “it remains possible that the previously observed difference between hypothetical WTP and actual contributions to public goods could largely be explained by free riding rather than be a reflection of upward bias in hypothetical answers,” (p. 106).

For private goods, Dickie, Fisher, and Gerking (1987) found no statistically significant difference in the hypothetical and actual prices for strawberries, although more recent research on these data (using alternative empirical estimation techniques) has produced mixed results (see footnote 1 in Poe, Clark, Rondeau, and Schulze, 2002). This is likely because private goods “do not suffer the free-riding problem that is seen with public goods” (Carson and Hanemann, 2005, p. 915).

The use of the CVM to estimate house values in contaminated areas will elicit reasonably valid value estimates. Houses are private goods that do not suffer from the free-rider problem, which can be a major source of bias; plus residents have an intimate knowledge of the homes in which they reside. Residents’ opinions expressed in surveys can therefore be well-informed valuations.\textsuperscript{4} Agarwal (2007) discusses homeowners’ estimates in comparison to lenders’ estimates of house value
(presumably purchase price updated via a housing price index) in a panel data set covering 82,000 homes. Mean absolute difference was 13.1% and mean values were overestimated by 3.1%. While significant, these homeowner-stated values seem to be reasonably accurate estimates, likely more precise than many jurisdiction’s tax assessment valuations. One should also keep in mind that the owner’s estimate could in some cases be more accurate than the bank’s estimate; so merely comparing differences does not tell us whose value is more likely to reflect market value.

Marketing and Survey Research Literature

Survey research methods have a long and less controversial history in market research and political polling. Companies, social science researchers, political candidates, and others routinely use surveys to gather information on a wide variety of topics, including prices. In the marketing research field, surveys may elicit WTP as part of demand forecasting for new products or willingness to pay for specific product characteristics.

Market researchers also use another stated preference method called “conjoint analysis,” that offers choices between alternative “packages” of various product characteristics to consumers. This allows for the evaluation of characteristics (such as an extra den or higher-quality kitchen) in the context of other property features (i.e., capturing interaction effects) and cost trade-offs. Another possible use of conjoint analysis is to determine respondents’ preferences for how far away they would like to live from contamination sources, such as power plants or zinc smelters.

Extensive use of survey methods in market research confirms that businesses find that surveys provide valuable information for business decisions. This literature is too extensive to be reviewed here; however, the market research literature provides a great deal of helpful advice on the use of survey methods to obtain value estimates that prove useful for firms’ product design, pricing, and the perceived quality of store brands (Sprott and Shimp, 2004). Survey research textbooks, such as Dillman (2007), cover survey research issues, including sample design, methods of data collection (phone, mail, Internet, and personal interview), question wording, survey instrument pre-testing, interviewer training and supervision, answer rotation, follow-ups and handling of refusals, and protocols for recording and analyzing data, etc. In summary, businesses rely on well-designed stated preference methods to elicit responses accurate enough to be used as a reasonable basis for business decisions.

Why are CV Methods Controversial?

Critiques from Economics

Normally transaction prices reflect a buyer who has sought full information, evaluated it by rational thought processes, considered competitive conditions, and bid without duress according to self-interest. Survey responses may or may not reflect the same
degree of care in gathering, processing, and acting on information to arrive at a valuation. However, both transaction prices and survey responses can be evaluated to determine how much emphasis to place in the reported valuation.

The CV literature concerned with inflated value responses emerges from the hypothetical nature of the exercise; this is called hypothetical or non-zero background bias (Schkade and Paine, 1994; Cummings, Harrison, and Ruström, 1995). Efforts to de-bias numerical survey responses trace back to Fischhoff (1982), who proposed to present respondents with a script that makes them aware of certain response effects and then to avoid them in their answers. Cummings and Taylor (1999) applied this strategy to a CV experiment to deter hypothetical bias and found that their script lowered WTP estimates. Bulte, Gerking, List, and de Zeeuw (2005) obtained the same result with an abbreviated script that simply suggests that the CV result might influence policy. Scripts also may curb inflated acceptances for higher bids (Brown, Ajzen, and Hrubes, 2003).

In some circumstances, however, responses to CV questions will tend to be conservative (Farmer and Lipscomb, 2008; Lipscomb and Koford, 2011). If respondents conjecture that a CV survey could impact a real policy result, this could be sufficient to induce conservative responses to an initial discrete choice question. The referendum frame, where respondents vote yes or no to a proposed policy change or bid amount, valued as a device to approximate the TILI condition, may not be the most robust vehicle in every political setting to establish this incentive to give an accurate response in all cases, especially for active policy issues.

Incentives in follow-up CV question experiments offer some guidance. Bohara et al. (1998) conducted a series of tests to argue that follow-up questions are deflated out of a concern to contain project costs. Cost concerns arise if follow-up questions signal the presence of bargaining, which violates the TILI condition to induce conservative responses. If the two expectations [that the results 1) might be used and 2) might be revised if rejected] exist prior to the follow-up, even prior to the survey, this would affect response incentives to a discrete choice (DC) question. This reinforces the idea to follow closely a real-world context.

The above discussion demonstrates that CV research has given considerable thought and investigation to circumstances that could result in biased responses. Reflection on these issues tells us that responses will be more accurate if respondents are knowledgeable or have direct experience concerning the value and characteristics of the good they are being asked about. For example, a New Jersey condominium shopper who has recently purchased a new condominium after extensive shopping would likely report a more accurate WTP estimate for an extra bathroom in a similar project than she would report regarding her WTP for preserving the water quality of a river in Tasmania.

A second major concern is whether or not respondents see any benefits in inflating or deflating their answers. The incentives facing respondents need not be direct, but could be indirect as well. If the respondent thinks of herself as an environmentally
friendly person, she might place an artificially high value on the water quality of that Tasmanian river. It is quite obvious that people can intentionally inflate or deflate estimates if they have a motive to do so, or even due to some behavioral economics bias common to most people in similar situations. These potential “self-interest” biases suggest that, when asking for damage estimates in contamination cases, more credence can be placed on the responses of control group property owners outside the affected areas (who have no incentive to inflate their damage estimates).

Critiques from Real Estate

Jackson (2004) and Roddewig and Frey (2006) reject the CVM because it is not one of the “three approaches to value” commonly used for real estate appraisals. Roddewig and Frey (2006) and Mathews (2008) denounce the CVM as a valid method because in particular cases CVM estimates did not match subsequent sales prices. Roddewig and Frey (2006) discuss four cases where CV survey responses overestimate the diminution of house values as evidenced by later transactions. They mention that part of the nature of the CVM is that it fails to incorporate many of the factors that go into purchase and sales decisions and does not capture trade-offs that are part of any economic decision due to opportunity costs, budget constraints, etc. Alternative explanations for the discrepancies, however, could be that the transaction prices reflect uninformed buyers or sellers (i.e., a form of market failure) or simply that market conditions changed between dates of transactions and when the survey data were collected. As noted, differences between revealed and stated preferences should not automatically be assumed to reflect “errors” in the latter. Both could be subject to human error or some form of market failure.

Matthews (2008) argues that hypothetical situations in CV surveys often fail to connect to real situations. Jackson (2004, p. 313), reinterpreting the work of Kinnard (1992), argues that “analysis and conclusions drawn from abstract techniques, such as contingent valuation surveys, could overstate any adverse impacts of environmental contamination on property values.” Wilson (2006) argues that the CVM should only be applied to non-economic goods. In many cases, such as groundwater contamination, pollution externalities can be associated with a private good (called internalizing the externality in the economics literature) in the same way that air quality levels can be attributed to properties at certain locations. Where markets are not fully informed about such issues, CVM estimates can be more indicative of value diminution due to contamination than transaction prices. A CV survey question can focus attention on the contamination itself, ceteris paribus, which allows consumers to make a more informed evaluation of its effects. Transaction prices may obscure effects of contamination through confounding with other variables or reflect uninformed decisions or faulty evaluations of risks. In some markets, the psychological defense mechanism called “denial” may cause people to undervalue health risks due to contamination. Some smokers deny the well-established health risks of cigarettes, displaying a similar psychological bias to suppress or ignore disturbing information. This is called disconfirmation bias and has been studied in the real estate literature by Seiler, Seiler, Traub, and Harrison (2008).
Avoiding Biases so that Survey Responses Approximate Market Values

Scholars note that where respondents may actually have the opportunity to purchase a good themselves, survey responses are more reliable. Where no or little market data are available, enough information must be provided to avoid the survey respondent filling in the gaps with their own information or conjectures. But, there is a trade-off to be recognized, specifically that too little information can lead to conjecture but too much information can lead to overload. Markets consist of buyers and sellers, as well as goods and their substitutes. By creating a scenario similar to the marketplace, were it in existence and functioning, the survey administrator can limit informational bias (Ajzen, Brown, and Rosenthal, 1996).

Respondents may bias their answers in WTP questions in a way they feel will please the sponsor or interviewer and that may not reflect their true WTP (Mitchell and Carson, 1988). For instance, if Coca-Cola were using a CV survey, respondents might answer in a way that was favorable to Coca-Cola, creating sponsor bias or funding bias. In like manner, the interviewer must pay particular attention not to indicate any preferred response indicators. This is why third-party survey vendors, who train their interviewers to read the scripts and questions verbatim and have no stake in the outcome, are useful.

Where the consequences of answers are perceived to have no effect on the respondent directly, hypothetical bias is said to occur. To mitigate this bias, survey questions must be asked in such a way that the respondents answer as if they are in a real market, or as if they will actually be affected by their decisions and not as if they are answering a question without any economic consequences. Cheap talk scripts can be effective at mitigating hypothetical bias (Cummings and Taylor, 1999). But, at least one study (Loomis, Gonzalez-Caban, and Gregory, 1994) finds that reminders of substitute goods and budget constraints have no influence on mean WTP estimates.

Mitchell and Carson (1989) provide a wide range of biases and potential threats to validity. Below is a summary of the biases that may impact WTP estimates, particularly in a real estate context.

- **Anchoring or Starting Point Bias**: Anchoring occurs when the elicitation method or payment vehicle introduces a WTP amount that influences the subsequent WTP amount(s) articulated by the survey respondent. This bias may be a tendency to ‘yea’ saying,” or saying ‘yes’ more often than one would expect. Alberini (2005) notes that the use of follow-up questions may cause respondents to anchor the value they place on the good under consideration to the bid amounts proposed to them in the survey questions. Anchoring results in WTP results that do not reflect a respondents actual WTP. This is one of the reasons researchers focus on dichotomous choice questions, where a single value is presented to respondents, who in turn simply answer ‘yes’ or ‘no’ to the valuation question. This also prevents respondents, especially in mail surveys, from reading ahead and anticipating the next question.
Range Bias: Range bias occurs when the survey presents a range of potential WTP amounts that influences a respondent’s stated WTP. For example, if we provide value ranges of $5–$10 or $10–$15 without pre-testing the question on respondents, we might influence responses upward or downward. The respondent’s true WTP or WTA maybe be below $5 or above $15. By introducing a range, one can introduce a bias that can easily be avoided (Whynes, Wolstenholme, and Frew, 2004). However, this illustrates a clear role for focus groups to test questions and valuation ranges prior to survey administration.

Relational Bias: Relational bias occurs when the description of the good to be valued presents information about its relationship to other public or private commodities (benchmark goods), which may influence a respondent’s stated WTP. In other words, if respondents focus on benchmark goods to assist in their valuation determination, their responses may be biased upward or downward depending on the price of the benchmark good relative to the good being valued. For example, in valuing a non-market good, this bias might be present if the description of the good being valued, say, water quality, includes the daily costs of treating water waste at the local waste treatment facility (which is likely already being paid for through one’s monthly water bill). A clear separation must be made in the survey so that respondents can distinguish between expenditures already being made related to water (e.g., monthly water bill) versus additional expenditures to be made solely for higher water quality.

Position Bias: The position/order in which valuation questions for different levels of a good (or different good) suggests to respondents how these levels should be valued. There are two issues to address here. First, the order in which questions are asked may influence the outcome. Second, the order in which the possible answers are provided (for closed-end questions) may influence the outcome. To eliminate the possibility of ordering effects, the researcher needs to specify that the third-party survey administrator randomize the order of certain blocks of questions. As an example, one person might be asked questions in this order: (1) What kind of impact did the paper mill have on your desire to purchase a home in this area? (2) What kind of impact did the power plant have on your desire to purchase a home in this area? (3) What kind of impact did the gypsum board plant have on your desire to purchase a home in this area? (4) What kind of impact did the food processing plant have on your desire to purchase a home in this area? Another person is asked the same questions, but in a different order: (1) What kind of impact did the food processing plant have on your desire to purchase a home in this area? (2) What kind of impact did the power plant have on your desire to purchase a home in this area? (3) What kind of impact did the paper mill have on your desire to purchase a home in this area? (4) What kind of impact did the gypsum board plant have on your desire to purchase a home in this area?
Then, to avoid ordering effects in the answers to particular questions, the survey administrator needs to randomize the order of the answers to certain questions (one person hears/reads these possible answers to a question: increases, does not change, decreases; another person hears/reads these possible answers to a question: decreases, does not change, increases).

- **Marginal Bidder Analysis:** Marginal bidder analysis is a tool to eliminate low bids and give higher bids more weight. Matthews (2008) argues that it is a poor mechanism. “In an actual transaction, the price that the buyer pays and a seller accepts typically results from the give-and-take between buyer and seller. Buyers and sellers both factor into their offers and acceptance the quality and quantity of available substitute homes, the presence of offers and other potential buyers, and their preferences,” (p. 264). This method eliminates extreme bids and gives more weight to those bids closer to the middle of the bid distribution. This is a form of trimmed mean analysis, which eliminates outlier observations. To present a balanced picture of the data, the real estate researcher should present the results of the marginal bidder analysis alongside other CV estimation methods common in the literature. A few of these include the WTP calculations following the Krinsky and Robb (1986) method, the Cameron (1988) method, and the Johansson (1995) method.

- **Vehicle Bias:** A respondent that is willing to pay a certain amount of their budget to obtain a certain service, may not be willing to have their taxes increased. Other types of bias may be included in surveys “where the method of payment may influence results” (Dickie, Fisher, and Gerking, 1987). In CV surveys that focus on estimating WTP related to environmental contamination, homeowners (the most commonly surveyed group) almost always have direct experience with mortgage payments, the most common payment vehicle in this situation. So, the familiarity with the payment vehicle reduces any bias that may arise. Also, the role of the cheap talk script (discussed below) is to emphasize that respondents should think in cash-equivalent terms, think of their budget constraints, and other relevant factors related to their statements of WTP.

Overall, this literature demonstrates that the CVM is well developed. A large number of biases have been identified, and ways to mitigate these biases have been vetted.

**MITIGATING CVM BIASES: A GUIDE FOR REAL ESTATE RESEARCHERS**

Researchers have suggested ways to mitigate the inherent biases of the CVM. This section presents ways to mitigate these biases in applied real estate work.
Survey Pre-testing

Wilson (2006) comments on the Arrow et al. (1993) NOAA Panel Report, a major study that established guidelines for the use of the CVM. Wilson emphasizes that pre-testing of the CV survey is one of the most important steps in collecting CV data. Pre-testing a survey can help survey administrators eliminate bias or other misinformation before a huge amount of effort has been expended.

Scenario Consistency

A key component of a CV survey is scenario consistency. Matthews (2008) notes that “Holding all else constant within the boundaries of a CV questionnaire requires special care. The specific information presented as a part of the hypothetical scenarios and CV questions must explicitly hold constant all features and events that can affect property values in both the baseline and contamination scenarios.”

Scope Test (Embedding Effects)

Scope tests relate to the idea that “respondents should be willing to pay more for a larger amount of a desired good,” (Carson, Flores, and Meade, 2001, p. 181). Diamond and Hausman (1994, p. 46) argue that insensitivity to scope (the embedding effect) is the “main contingent valuation anomaly.” Scope tests are easily conducted when the researcher requires all survey respondents to value different quantities of the good in question or requires different respondents to value different quantities of the good. Carson and Mitchell (1993) believe insensitivity to scope can be avoided with survey design and administration, thereby improving the validity of CV valuations. One way to do this is to present a series of options (e.g., how much are you WTP to preserve 2,000 birds) and then remind respondents of their previous answers as they answer subsequent valuation questions (e.g., much are you WTP to preserve 20,000 birds). One possible problem with this approach is that it might result in anchoring bias (where respondents anchor future responses to responses they have made in the past). Another possibility is to present each respondent with a unique number of birds to preserve and then test for a scope effect across respondents in aggregate. This might avoid the anchoring bias mentioned in the other option.

Question Wording

The book by Dillman (2007) contains an entire chapter on questionnaire design and how two different versions of a question can produce different responses. One of the most common mistakes is asking a question that contains a second embedded question (called a double-barreled question) or even worse a third embedded question (triple-barreled question). An example of a double-barreled question is: Do you live on a farm and are you actively engaged in farming? Clearly, answering ‘yes’ to this question suggests that a survey respondent both lives on a farm and is actively engaged in farming. To avoid the confusion, this question should be broken out into a two-part question:
Q: Do you live on a farm? YES NO
Q: If you answered YES above, are you actively engaged in farming? YES NO

Another example of a double-barreled question, from Bryman and Bell (2007), is:
How satisfied are you with your salary and job conditions? Once again, to avoid confusion, this question needs to be broken out into two different questions.

Certainty Statements

Certainty statements have become an important validity check on WTP estimates. These statements are designed to elicit the respondent’s certainty about WTP to identify respondents who would actually do what they say. Certainty typically is measured using one of two scales: (1) a Likert-style 10-point scale, with higher numbers indicating more certainty, or (2) a “probably sure, definitely sure” scale. Below we provide examples of the two certainty statement questions.

Q: How sure are you that you would pay $150,000 for a residence you were interested in if it were located within one-half mile of (the environmental contamination)? Would you say you are...?
1. Definitely sure
2. Probably sure
3. Not very sure
4. Not at all sure

Q: On a scale from 1 to 10, how certain are you of your answer? One (1) indicates you are very uncertain; 10 indicates you are very certain of your answer.

<table>
<thead>
<tr>
<th>VERY UNCERTAIN</th>
<th>VERY CERTAIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

The certainty statements are used to adjust hypothetical willingness to pay responses based on the level of certainty elicited, the idea being that respondents with a high level of certainty are more representative of individuals making an actual purchase (Champ, Bishop, Brown, and McCollum, 1997). Respondents who say they are definitely sure they would pay are more likely to behave in a way that is consistent with their stated intention. Econometrically, certainty statements have been used to calculate WTP estimates for those respondents who are more certain that they would actually pay the amount agreed to in the referendum question (e.g., Blomquist, Blumenschein, and Johannesson, 2009). Certainty statement variables have not been used as independent variables in regression analyses.

Cheap Talk Scripts

Cheap talk scripts were designed to mitigate hypothetical bias by reminding survey respondents about their budget constraints and to reinforce the idea that a real cash payment could be involved. There is some debate about the length of cheap talk scripts
Before we have our vote, I want to talk to you about a problem that we have in studies like this one. As I told you a minute ago, this is a hypothetical referendum—not a real one. No one will actually pay money at the end of the vote. But I also asked you to respond to the vote as though the result of your vote could involve a real cash payment by you.

And that’s the problem. In most studies of this kind, people seem to have a hard time doing this. They vote differently in a hypothetical referendum, where they don’t really have to pay money, than they do in a real referendum, where they really could have to pay money. For example, in a recent study, several different groups of people voted on a referendum just like the one you are about to vote on. Payment was hypothetical for these groups, as it will be for you. No one had to pay money if the referendum passed. The results of these studies were that on average, across the groups, 45% of them voted YES. With another set of groups with similar people voting on the same referendum as you will vote on here, but where payment was real and people really did have to pay money if the referendum passed, the results on average, across the groups were that 27% voted YES. That’s quite a difference, isn’t it?

We call this a “hypothetical bias”. “Hypothetical bias” is the difference that we continually see in the way people respond to hypothetical referenda as compared to real referenda. In the real referendum, where people knew they would have to pay money if the referendum passed, 27% voted yes and 73% voted no. When payment was hypothetical and people knew they would not pay anything if the referendum passed, just like your vote today, 45% voted yes and 55% voted no.

How can we get people to think about their vote in a hypothetical referendum like they think in a real referendum, where if enough people vote YES, they’ll really have to pay money? How do we get them to think about what it means to really dig into their pocket and pay money, if in fact they really aren’t going to have to do it?

Let me tell you why I think that we continually see this hypothetical bias, why people behave differently in a hypothetical referendum than they do when the referendum is real. I think that when we hear about a referendum that involves doing something that is basically good—helping people, improving environmental quality, or anything else—our basic reaction in a hypothetical situation is to think: sure, I would do this. I really would vote YES to spend the money—I really, really, think I would. What our YES vote really means in this case is that we are basically good people, and that we would like to see good things happen.
But when the referendum is real, and we would actually have to spend our money if it passes, we think a different way. We basically still would like to see good things happen, but when we are faced with the possibility of having to spend money, we think about our options: if I spend money on this, that’s money I don’t have to spend on other things. If I spend money to help needy students, that’s money I don’t have to spend on groceries, go to a movie, or perhaps give to some environmental organization. So when the payment is real if the referendum passes, we vote in a way that takes into account the limited amount of money we have. We vote realizing that we just don’t have enough money to do everything we might like to do. This is just an opinion, of course, but it’s what may be going on in hypothetical referenda.

In any case, the only way that we know to get people like you to vote in our hypothetical referendum just like you’d vote if the referendum was real is to simply ask you: in the vote that we’re going to take in a few minutes, please do the following:

- Think about what you’re voting on. If this were real—if more than 50% of you voted YES, you would actually have to dig into your pocket and pay $8.00 right now—do you really want to support the UMass–Amherst Scholarship Fund enough that you would be willing to spend the money?
- Also, let me make clear that the $10.00 participation fee that you were paid today is your money. You’ve spent your time helping us in our research, and you’ve earned it! You were told that the money is yours, believe it! So, if I were in your shoes, and I was asked to vote YES or NO on this proposition that requires all of us to pay $8.00, I would think about how I feel about spending my money this way. When I got ready to vote, I would ask myself: if this were a real referendum, and I had to pay $8.00 if the referendum passed, do I really want to spend my money this way. If I really did, I would vote YES; if I didn’t, I would vote NO—I wouldn’t throw my money around. That’s just my opinion, of course. You must do whatever you want to do.
- In any case, I ask you to vote just exactly as you would vote if you were really going to face the consequences of your vote: which is to pay money if the proposition passes.

Despite the depth of coverage by such a long cheap talk script, there is some concern about information overload and “daydreaming.” The Cummings and Taylor (1999) script, like the one presented here, is long (eight paragraphs) and would thus be difficult to use in a telephone or mail survey. A much shorter script, such as that used by Loomis, Brown, Lucero, and Peterson (1996), might reduce hypothetical bias while accommodating respondents’ attention spans.
Exhibit 1
“Central Tendency” and Ranges of Values

Survey responses indicating value

Transaction prices indicating value

UNCERTAINTY IN VALUATIONS AND TRANSACTION PRICES

Both transaction prices and survey responses stating WTP are human decisions and are made with degrees of uncertainty. In some circumstances, we may observe upward or downward bias and/or differences from subsequently revealed ex post fundamental values. When survey responses and transaction prices disagree, one should not automatically conclude that the differences are due to errors in the survey values. Transaction prices could also be far from market values under particular market conditions.

Kummerow (2000, 2002) argued that before a transaction occurs there is not one possible price, but rather a distribution of possible prices with some prices more probable than others. Each observed sale price or stated value is therefore a “draw” or “event” from a distribution of possible values. There is no “true” value in the sense of a measurable quantity like height or weight. This is one reason why market value is generally defined as “the most probable price” and why appraisers provide “opinions of value.”

Exhibit 1 depicts the valuations from revealed preferences (arms-length sales) and stated preferences (survey responses) as a range of possible values relative to a “central tendency” of a value distribution. Note that both potential transactions and survey responses reflect a range of possible prices for each property, in part reflecting realistic assessment of the potential range of prices given limited information, negotiations, circumstances of sale, and other factors. In Exhibit 1, the wider range of survey responses reflects the assumption that there may be more uncertainty in survey responses for the reasons stated above. It is important to note that a transaction price in a particular instance could be more biased, while a particular survey response could be more accurate.

At the heart of this uncertainty discussion is the efficient market hypothesis (EMH), which generally was accepted as conventional wisdom from roughly 1970 to 1990. Several journals address the irrational or behavioral aspects of market pricing, including the Journal of Economic Psychology, Journal of Behavioral Finance, and Journal of Economic Behavior and Organization. Also, Sewell (2008) posted a
“History of the Efficient Market Hypothesis” list of about 140 references that describe the evolution of the EMH concept, critiques of the EMH, and empirical results from behavioral experiments.

Behavioral economics experiments and market experience demonstrate irrational pricing decisions under particular circumstances due to the limitations of human information processing capacity and biases in the way our minds construct valuations. People offer different valuations depending on how the issue is presented to them (e.g., Genesove and Mayer, 2001; Clauretie and Thistle, 2007; and Einio, Kausti, and Puttonen, 2008; for commercial real estate, see Bokhari and Geltner, forthcoming). These framing issues could affect transaction prices, as well as values elicited by the CVM.

If the EMH does not hold, then revealed preference methods may lose some of their strength and stated preference methods may become more relevant as a tool for determining value. It might even be the case that the CVM, based on the presentation of more balanced information, could give more reliable value estimates than inferences from transactions in markets that are not in equilibrium.

**USPAP**

The Uniform Standards of Professional Appraisal Practice (USPAP) do not speak directly to the CVM, but it does offer some indirect guidance about the role it might play in valuation assignments. A few relevant intersections between USPAP and the CVM are highlighted here.

**Scope of Work Rule**

Regarding the Scope of Work Rule, USPAP (2010–2011, p. U-13) states that: “Appraisers have broad flexibility and significant responsibility in determining the appropriate scope of work for an appraisal, appraisal review, and appraisal consulting assignment. Credible assignment results require support by relevant evidence and logic. The credibility of assignment results is always measured by the context of the intended use.”

In markets that are in disequilibrium or experiencing some kind of market failure, it is reasonable to argue that sales transaction prices do not necessarily reflect market values. In situations like this, collecting additional survey data and logically connecting survey data results back to the relevant market under duress increases the credibility of an appraiser’s assignment results.

**Competency Rule**

Regarding the Competency Rule, USPAP (2010–2011, p. U-11) states that: “The appraiser must determine, prior to accepting an assignment, that he or she can perform the assignment competently. Competency requires: 1. the ability to properly identify the problem to be addressed; 2. the knowledge and experience to complete the
assignment competently; and 3. recognition of, and compliance with, laws and regulations that apply to the appraiser or to the assignment.”

Competency in an appraisal assignment requires that the appraiser disclose the incompetency at the beginning of the assignment and discuss how she will achieve competency prior to the completion of the assignment. Competency could be achieved through individual study, attending a seminar on the particular subject, or other ways. The CVM particularly, and survey research generally, is one of the more advanced subjects that an appraiser may face. From questionnaire design to non-linear regression models used to estimate WTP, achieving competency in the use of the CVM is very cumbersome.

Advisory Opinion 9

USPAP (2010–2011, p. A-16) Advisory Opinion 9 specifically covers the topic of real property valuation in the presence of environmental contamination. Of interest regarding the CVM is the discussion of valuation issues for impaired properties. Standards Rule 1-4 states that: “In developing a real property appraisal, an appraiser must collect, verify, and analyze all information necessary for credible assignment results.” The discussion points out that diminution may not be calculated simply by measuring remediation costs. Rather, value may be affected by cost, use, and risk effects. Specifically, risk effects are of interest here.

“Risk effects are typically estimated by the appraiser and often represent the most challenging part of the appraisal assignment. These effects are derived from the market’s perception of increased environmental risk and uncertainty. The analysis of the effects of increased environmental risk and uncertainty on property value (environmental stigma) must be based on market data, rather than unsupported opinion or judgment.

In general, the unimpaired value of the property being appraised can be estimated using the sales comparison approach [SR 1-4(a)], cost approach [SR 1-4(b)], and income approach [SR 1-4(c)]. Estimating the effects of environmental contamination on real property value usually involves the application of one or more specialized valuation methods. These methods should be consistent with the requirements related to the valuation approaches in USPAP,” (p. A-19).

“The market’s perception” could be measured by the CVM or other related methods, such as conjoint analysis. In short, we interpret the USPAP to allow CVM, consistent with the many public agencies that use CVM for valuing public goods. This would be particularly appropriate for situations involving contaminated property, as long as it is applied in a well-reasoned and supportable way. For example, market transactions data (the sales comparison approach) could be complemented with survey responses, as has been done by Earnhart (2001, 2002). Furthermore, Advisory Opinion 9 encourages the gathering of data that provide insight into the risks surrounding contaminated properties; this is exactly the function and intention of the CVM as we have articulated in this research.
CONCLUSION

The CVM derives estimates from different data than the three standard approaches to valuation and can add more depth to market analysis and valuation conclusions. CVM is an approach that can be used to complement, correct, or externally validate opinions of value from the three standard approaches.

Valuation is ultimately an exercise in modeling economic behavior. Market transactions are determined through the interaction of buyers and sellers; and this produces a measure of value under a given set of conditions. This measure of value remains an opinion, subject to error and uncertainty, limited information, and bounded rationality. Within this context, there is nothing particularly troubling or unique about survey responses as indicators of value. Survey responses are stated opinions that might, under some circumstances, be better informed and more utility maximizing than less informed revealed preferences.

The strengths of survey methods offset the weaknesses of transaction-based models, especially when “frozen” markets produce very few sales or when market transactions do not include symmetric information disclosure. A transaction gives a value reflecting real purchase behavior, but the value is the only datum. Use of the CVM allows a deeper exploration of how market participants came to their value conclusions. When values derived from the CVM differ from transaction prices, both should be examined to determine if there are sources of bias that can be corrected. Suspicion of market failure makes the CVM an even more applicable approach when working with contaminated property.

ENDNOTES

1. Tracing back to the seminal article by Akerlof (1970), it is well established that asymmetric information affects economic decisions. Information allows individuals to make choices that yield higher expected payoffs or expected utility than could be made in the absence of that information.

2. Existence values refer to the value individuals place on things like parks they will never visit or, for example, charismatic species such as grizzly bears or Siberian tigers that they hope never to meet in person.

3. These tests assume that actual WTP measures true compensating variation, which is the amount of additional money a person would need to reach her initial utility level after a change in the prices of goods, or a change in the quality of the good, or the introduction of new goods (Hicks, 1939).

4. The U.S. Census Bureau and some tax assessors rely on survey methods involving self-reports of home values, which suggest that values originating from surveys are reliable enough for empirical work. Given that transactions data in contaminated areas are often scarce or reflect some kind of market failure, the CVM can provide valuable additional market value information.
REFERENCES


International Association for Assessing Officers. Standards on the Valuation of Properties Affected by Environmental Contamination.


The authors would like to thank Brandon Koford, David Keating, Ben Hoen, two anonymous referees, and participants at the 2010 American Real Estate Society Annual Meeting for their helpful comments.

Clifford A. Lipscomb, Greenfield Advisors, LLC, Atlanta, GA 30339 or cliff@greenfieldadvisors.com.

Max Kummerow, Greenfield Advisors, LLC, Seattle, WA 98121 or max@greenfieldadvisors.com.

Will Spiess, Greenfield Advisors, LLC, Seattle, WA 98121 or will@greenfieldadvisors.com.

Sarah Kilpatrick, Greenfield Advisors, LLC, Seattle, WA 98121 or sarah@greenfieldadvisors.com.

John A. Kilpatrick, Greenfield Advisors, LLC, Seattle, WA 98121 or john@greenfieldadvisors.com.