Constructing markets: environmental economics and the contingent valuation controversy

Banzhaf, H. Spencer

Department of Economics Andrew Young School of Policy Studies
Georgia State University

October 2016

CONSTRUCTING MARKETS:

ENVIRONMENTAL ECONOMICS AND THE CONTINGENT VALUATION CONTROVERSY[*]

H. Spencer Banzhaf[†]

Georgia State University, PERC, NBER

October 2016

I. Introduction

The history of applied economics in the mid-to-late 20th Century involved applying increasingly abstract economic constructs to increasingly intangible objects. For example, in consumer theory, economists moved from measuring value with market prices to using various forms of consumer surplus. At the same time, from measuring consumers' values for the quantity of material goods, they moved on to value the quality of those goods and the value of intangible goods not traded in markets, like the preservation of the natural environment. In doing so, they faced a new challenge: How could they observe such values when there were no markets and so no prices to observe first?

One seemingly simple answer to that question was to ask people through surveys. Over

[*] I thank Roger Backhouse, Béatrice Cherrier, Steve Medema, Brian Murray, two anonymous referees, and participants in the 2016 HOPE conference for helpful comments.

[†] Dept. of Economics, Georgia State University, PO Box 3992, Atlanta, GA 30302. hsbanzhaf@gsu.edu.

the years, a broad cross section of economists had suggested doing just that, including S.V. Ciriacy-Wantrup (1947) and Irving Fox[1] (two resource economists), Arthur Bowley (1920) (to see how welfare changed with of cost-of-living indices), Thomas Schelling (1968) and Howard Raiffa (1969) (on health risks), and, perhaps in a weak moment, even George Stigler (1943). But actual applications of what became known as the contingent valuation (CV) method really took flight in the 1960s and early '70s with applications to valuing environmental amenities.

The CV method became especially important when, around the same time, environmental economists introduced the notion of "existence values", i.e., the value of preserving a resource simply for its existence, regardless of how people might interact with it behaviorally (Krutilla 1967), thus taking the abstract and intangible to a new level. The CV method was crucial to applications to measuring existence values because, practically by definition, existence values involved no behavioral actions that could be observed in the world outside a survey.

Nevertheless, the CV method faced an uphill battle for acceptance within the wider intellectual community of academic economics. Notwithstanding occasional suggestions to use them, the mainstream of the profession had long been skeptical of surveys. For example, Milton Friedman and others strongly opposed them, both as used by Richard Lester to interview businessmen about their behavior (Boulier and Goldfarb 1998, McCloskey 1983) and as used by L.L. Thurstone to estimate utility functions over consumer goods (Moscati 2007). Friedman argued on methodological grounds that individuals' own understanding of their actions need not match a scientific description of them. Additionally, Samuelson (1954) had argued that selfish agents would give deceptive signals about their value for public goods to influence the outcome.

[1] Memorandum from Irving K. Fox, Dec. 18, 1950, Box 2227 (River Basin Studies—Economic & Recreational Benefits, 1949-1953), Entry 11, Record Group 79 (National Park Service), National Archives Building, Washington, DC.

Though opposition to surveys was by no means unanimous, if Friedman and Samuelson agreed to team against something, it was bound to face tough sledding.

Beyond skepticism of surveys generally, CV faced a more specific barrier to widespread acceptance. Precisely because it measures such abstract entities, a CV survey does not merely elicit opinions or record concrete behaviors. It creates a hypothetical market for goods not normally traded in markets, asks respondents to think about their behavior "contingent" on the market being real (behavior which generates bids), and then observes the generated bids. Thus, unlike simpler surveys, it is not merely a recording device documenting relatively objective behavior. Too, unlike statistically estimating a demand curve, the CV method does not involve observing behavior or using data that existed prior to the creative work of the economist. To the contrary, as with experimental economics, which developed over the same period from similar motives, CV is built around the idea of constructing observations.

Historians of economics now are used to the idea of economists constructing their observations (Boumans 2005, Maas and Morgan 2012) but economists are not, or certainly were not in the 1970s or '80s. Thus, CV put economists in an unfamiliar and uncomfortable position. Borrowing language from Callon (1998) and MacKenzie (2007), to the economists involved CV was all too obviously "performative"—meaning it does not just measure something, it does something in the economy. Indeed, setting aside some broad interpretations of performativity, CV meets J.L. Austin's original linguistic meaning: the respondent to a CV survey literally *speaks* economic values into being.

The natural question then becomes, just what has been spoken into being? Do the objects look like the kinds of values that would be generated by *homo economicus*? Economists began asking such questions early on, sometimes out of skepticism of the method, sometimes out of a

sympathetic effort to calibrate the market institutions to make sure they had the right one in place. A CV survey judged to be successful would generate behavior consistent with microeconomic theory of behavior, with consistent preference orderings, sensitivity to prices, and modest sensitivity to income. Various anomalies in the pattern of stated values, however, suggested it was more difficult to use CV surveys to fabricate values from genuine *homo economicus* than environmental economists originally had hoped. Over the course of the 1980s, these anomalies attracted increasing attention from the wider profession, and environmental economists soon found themselves enmeshed in broader debates about experimental economics and behavioral economics.

To generalize, environmental economists had three options open to them when interpreting these anomalies: (i) CV potentially could generate data consistent with *homo economicus* once best practices were determined, while in the meantime poor CV practices occasionally generated the anomalies; (ii) CV could be inherently flawed; or (iii) CV could be generating the right data, so the fault in the anomalies lay not with the procedure but with the economic model of behavior. Behavioral economists like Daniel Kahneman, who teamed up during the 1980s with Jack Knetsch, an environmental economist working with CV almost from the beginning, argued for the latter position. Once CV appeared to support such criticism, skeptical economists who might otherwise have ignored CV were motivated to attack.

Meanwhile, the stakes were increasing. First, a series of regulatory actions gave federal agencies more incentive to apply CV to a wider range of policy settings, thus building political support for the method. Then, following the 1989 *Exxon Valdez* oil spill, Exxon funded a large literature critiquing CV, thus generating an opposing political force. Economists on all sides of

the debate wrote influential review articles, but in the halls of mainstream economics departments, CV became decidedly unfashionable following these debates. Nevertheless, it continues to be used by academic economists in other parts of the university (like agricultural economics or health policy departments), government economists, and other applied researchers, with over 7,500 studies and papers from over 130 countries collected in one recent bibliography (Carson 2011).

This paper tells this story in five parts. Section II introduces the first CV studies in the 1960s and the policy context in which they were born. Section III brings us to the growing literature in the 1970s and '80s testing the validity of the CV method, and the anomalies researchers found and debated. In Section IV, those debates come to a controversial head following the 1989 Exxon Valdez oil spill. Finally, Section V offers a concluding postscript to the CV controversy.

II. Robert K. Davis and the "Interview Method"

The basic problem for environmental economics of valuing non-market goods grew out of benefit-cost analyses of dams and other water projects (Banzhaf 2009, 2010, Hanemann 1992). In the late 1940s, various political pressures combined to force applied economists working in this setting to incorporate the value of outdoor recreation into the benefit-cost calculations—quite against their better judgment (Banzhaf 2010). Canvassed for ideas on how recreation might be valued, most economists said it could not be done: recreation was too spiritual, too intangible to be monetized. However, in the late 1950s and early '60s, Marion Clawson and Jack Knetsch, two Harvard-trained economists working at the new Ford-supported think-tank Resources for the Future (RFF), developed an approach to the problem based on the idea that the observed trips people take to a recreation site and the distance they are willing to travel to it could provide the

information needed to back out a demand curve, an approach known as the travel cost method.[2] Wrestling with what sort of theoretical construct in value theory actually could be measured with this new tool, they came to the view that "consumer surplus," was the appropriate construct, despite concerns that no real-world market institutions gathered such surplus as revenue (Banzhaf 2010).[3]

This earlier experience valuing environmental amenities provides three salient insights that extend into the story of CV. First, economists were wrestling with the problem of applying an economics still defined as a science of material welfare to intangible or spiritual settings like outdoor recreation. Second, theory does not always precede measurement; sometimes so-called "applied" work can hardly be said to apply a clear theory, but nevertheless operates within an intellectual structure (like benefit-cost analysis in this case) that requires quantification. In these cases, measurement can give rise to new theories about what has been measured (or, in the case of consumer surplus, new acceptance of old ones). And finally, whether a constructed economic value could be captured by a particular market institution and whether that institution fit a particular historical and policy context remained an important consideration well into the supposedly institution-free neoclassical era.[4]

The first application of what would come to be called CV was proposed as an alternative

[2] In their recent essay on applied economics, Backhouse and Cherrier (2016) consider the importance of the shifting sites where economic knowledge is produced, from academia to think tanks and government agencies. In the case of non-market valuation, it is notable that the istory repeatedly goes through Harvard University and two Ford-supported think-tanks, the RAND Corporation and RFF, as well as through agricultural economics departments and government agencies (Banzhaf 2009, 2010, 2014).

[3] Consumer surplus is a construct in microeconomic theory capturing the additional value consumers have for a good beyond what they actually have to pay.

[4] Banzhaf (2010) further explores these three points.

to the travel cost model for valuing outdoor recreation. For his dissertation in economics at Harvard University, Robert K. Davis estimated the economic value of outdoor recreation in the Maine woods using what he called "the interview method" (Davis 1963a,b). Davis had received a BS and MS in agronomy from Ohio State and worked at Harvard under Otto Eckstein, with assistance from John Kenneth Galbraith and Edward Banfield (a political scientist). According to Carson (2011), Davis was unaware of earlier suggestions to use surveys and developed his ideas while taking a course from Stanley Stouffer, one of the leading US survey researchers. Davis was supported by an RFF dissertation fellowship and completed his dissertation there. While at RFF, he worked with Knetsch to compare the CV method with the travel cost method applied to the same resource, finding similar results (Knetsch and Davis 1966).

Davis's work followed an outline that would be used in most CV surveys going forward. CV surveys since that time have tended to comprise four basic steps. First, ask "warm up" questions about the respondents' general attitudes toward the subject matter, to get the respondent in the right frame of mind. Second, provide information about the specific policy context to be valued. Third, at the heart of the survey, construct the market mechanism for "eliciting" values (as some would have it) or "constructing" them (as others would put it). This market has two crucial components, the "payment vehicle," which is the hypothetical channel through which households would make their payment (e.g., taxes, user fees, etc.), and the "elicitation question," which is really the rules of behaving in the market (e.g., stating a value, submitting a "bid," or accepting/rejecting a set cost). Fourth and finally, ask follow-up questions about how respondents perceived the survey and about their income and other demographic characteristics. (See Mitchell and Carson 1989 for a state-of-the-art overview as of the period covered here).

From a historical perspective, we also can see in Davis's work three themes emerging at

the outset of the CV literature that would set the context for much of the future debate. First, CV applications were, from the beginning, applied to measuring the benefits of intangible services, like outdoor recreation, which economists until that time had largely said could not be quantified and indeed were not even economic. Previous applications to outdoor recreation using travel costs were still quite new and controversial. Consequently, in his dissertation, Davis felt the need to directly address the idea that recreation belonged to the realm of the aesthetic rather than the economic, arguing there should be no distinction: many market goods have aesthetic dimensions to them.

Second, unlike other early work in the valuation of outdoor recreation up to that time, Davis showed no hesitation about the relevant construct in economic theory that he wished to measure: Hicksian compensating variation, or the most that a consumer would be willing to pay for a resource and be no worse off than he would be without it ("willingness to pay" [WTP], for short). In doing so, Davis tied CV to a particular economic theory of individual welfare, one which, in principle, carried with it a set of testable hypotheses. However, the theory entailed very abstract constructs that, to most economists at the time, were unobservable. As Hicks had shown, WTP could be viewed as a variant of consumer surplus, in which income is adjusted along the demand curve in such a way as to hold utility constant. However, since such adjustments are never made in the real world, such demand curves seemed to most to be highly artificial. To overcome this problem, Davis created an artificial world in his survey where such entities could be measured. Thus, the survey involved two departures from the standard practices of the time: a highly abstract construct like WTP was being applied to a highly intangible object like outdoor recreation.

Third, to Davis and others early workers on problems of nonmarket valuation, the market

institution that would be created was important in itself. Davis's over-arching strategy was to create an "approximation of the market" (Davis 1962a). But that strategy only begged the question of what market. What institution, with what hypothetical allocation rules, should he create? Davis employed what he called a "bidding game." Interviewing people who already visited recreation facilities in the Maine woods, he asked respondents to consider their recreation experience and whether they would still have come if their transportation costs for reaching the destination were $X higher. If they answered "yes," Davis raised the cost by a set increment until they said no; if they answered "no" originally, he lowered the cost until they said yes. In this way, he converged on their maximum WTP. As Davis recognized, he was essentially asking for an all-or-nothing deal, in which respondents could come or not come but not vary the length of their stay in response to their added cost. From the distribution of responses and the distribution of days-per-trip, he then constructed a demand curve for visitation days. Finally, he computed the consumer surplus under this demand curve.

A number of environmental economists soon picked up the CV method. Particularly influential was a series of studies by Alan Randall, a resource economist working during this period at the University of New Mexico and University of Kentucky. Randall's work represents a continuation and extension of these three themes in ways that would prove significant for the future history of CV.

First, Randall, Ives, and Eastman (1974) applied CV to an even more abstract commodity than had previously been tackled: the aesthetic benefits associated with cleaner air, including improved views of the landscape.[5] They used photographs to communicate the aesthetic levels

[5] Around the same time, Cicchetti and Smith (1973) similarly evaluated WTP for the *quality* of the recreation experience.

that would be obtained under different policies regulating a power plant. These aesthetic benefits were an example of "existence" or "nonuse" values—reasons for valuing the natural environment that have nothing to do with how it might be used (for outdoor recreation say), but simply because people appreciated them for their beauty or uniqueness. Such nonuse values were a new category of benefits first discussed in economic terms by Krutilla (1967). In this way, the intangibility that first confronted resource economics was now being taken to a whole new level (Banzhaf 2016). Importantly for this story, this move increased the importance of CV in the policy context because, as Randall et al. understood, contingent valuation was the *only* possible tool available for measuring such nonuse values. Whereas other methods for measuring environmental values relied on observing actual behavior (e.g., trips in the travel cost model), for nonuse values, there is no such connection to observable behavior. For example, one might value the existence of, say, a pristine Alaskan wilderness without ever visiting it. In this way, the objects created by CV surveys became even more divorced from observable behavior; they could only exist in the contingent world constructed by the survey.

Additionally, Randall extended the connection between CV and the microeconomic theory of individual welfare, while simultaneously emphasizing the importance of the institutions created in the hypothetical world of the survey. Randall's work is unique in combining an interest in CV with work on markets, transaction costs, and property rights, with a particular focus on the Coase theorem (e.g. Randall 1972).[6] This combination of interests was important for two reasons. First, it meant that Randall and his colleagues were particularly sensitive to the design of markets and the way people operated in them. Second, from his earliest work on the Coase theorem, Randall was thinking about how initial property rights allocations could affect market

[6] See Medema (2014) on Randall's role in bringing the Coase theorem to environmental economics.

prices and market transactions (Randall 1972, Randall and Stoll 1980, Brookshire, Randall, and Stoll 1980). If households lack the commodity initially, in the hypothetical world of the CV survey it would be natural to ask them to buy it, so the survey should elicit their WTP. If they do have a property right to the commodity, it would be more natural for them to sell it in a market setting, and so the survey should elicit the lowest willingness to accept (WTA) money in exchange. Microeconomic theories predicted that these differences should be small. Nevertheless, as we shall see, WTP/WTA comparisons became a major source of controversy for contingent valuation.

More generally, Randall's body of work focused attention on how to create what Randall, Ives, and Eastman (1974) called a "contingent market" (apparently thereby coining the term "contingent valuation"). Randall et al. argued that to do this successfully, the hypothetical situations had to be "realistic and credible" and "concrete rather than symbolic." Second, the contingent market should involve "institutionalized or routinized behavior, where role expectations of respondents are well defined" (136), a criterion they thought was met by the bidding game. Third, the payment vehicle had to be realistic and relevant to the respondents. In alternative versions of the survey, they thus used either sales taxes (realistic, as people are used to paying sales taxes to finance public goods, and relevant, especially to tourists) or alternatively higher electricity bills (realistic, as the cost of controlling air pollution feasibly could be passed onto customers, and relevant to local residents).

This over-riding concern for constructing a realistic contingent market was echoed in the first major effort to review and standardize best CV practices. Cummings, Brookshire, and Schulze (1986) wrote:

[F]or most pure public goods … market institutions do not exist. The CVM

> [contingent valuation method] is then used as a substitute for the "missing" market; it is used to simulate the market in the sense of eliciting revelations of preferences (a willingness to pay) analogous to those which would have resulted under market conditions. Like the market institution, the CVM must then be viewed as an "institution." *Thus, the general criterion against which to assess the CVM becomes clear: the extent to which the CVM institution, and preference revelations drawn therein, approximates the market institution and values derived therein*. (72, emphasis added)

In other words, the task of CV is to create a (hypothetical) market in which, in Randall et al.'s words, behavior was "institutionalized or routinized" and "where role expectations of respondents are well defined." Those expectations, of course, were to behave like *homo economicus*. CV studies could then be judged by whether the CV institutions mimic markets and further whether people behaved in them the way economic theory suggested they should.

III. Testing Contingent Valuation, Testing Economic Man

In his study of experimental economics, Guala (2007) distinguishes between experiments designed for building institutions and those for testing theories. In his view, experimental economics, or at least that branch of it which has been most successful, worked by institution building, and thus "by creating *homines economici*, not by questioning their existence" (Guala 143). Guala quotes Charles Plott:

> "The task is to find a system of institutions—the rules for individual expression, information transmittal, and social choice—a "process" that mirrors the behavioral features of the mechanism. The theory suggests the existence of processes that perform in certain (desirable) ways, and the task is to find them. This is a

pure form of institutional engineering." (Plott 1981, cit. Guala 2007 147)

While the goal was not to test the existence of economic man, nevertheless whether the institutions built in the lab had successfully been created remains an empirical question requiring testing.

That there is be a parallel here between CV and experimental economics should not be surprising. CV and experimental economics grew up together during the 1970s with common motivations: at their heart, both are concerned with understanding behavior that cannot be directly observed in markets, often surrounding questions related to public goods.[7] Some first generation experimental economists such as Vernon Smith and James Cox began their careers as natural resource economists. Indeed, many economists stood at the boundary between CV and experimental economics, including David Brookshire, Don Coursey, Ronald Cummings, Glenn Harrison, and William Schulze. The intersection was attractive, because laboratory experiments became one way of testing whether the hypothetical market institutions created in a CV survey successfully created *homo economicus*. Too, experimental economists like Vernon Smith and Charles Plott were drawn into debates over CV, which was something of a rival endeavor.

While both experimental economics and CV seek to create market institutions that reveal demand, differences in their ultimate objectives motivate concomitant differences in the institutions they construct. Experimental economics often begins by inducing preferences with the incentives of material rewards and then proceeds to test how different institutions perform at processing those incentives and producing desirable (rational and efficient) outcomes, which V. Smith (1982) called the "Hayekian hypothesis" (Guala 2007, Muniesa and Callon 2007). Thus,

[7] On the history of experimental economics, see Guala (2005, 2007), Moscati (2007), and Svorenčík (2015).

the institution (or machine) does the work. The job of experimental economists is to build the machines and to find the ones that produce desirable outcomes and that can be applied to real-world problems.

In contrast to this Hayekian approach (broadly construed), the CV method takes a regulatory approach. It is predicated on the pre-existence of a regulatory institution such as benefit-cost analysis (or, perhaps, a Pigouvian tax) already performing in the real world. The objective of CV then is to build a kind of temporary machine that can measure (or construct) WTP values that can be plugged into the benefit-cost analysis. That is, while the experimental economist produces machines that can themselves be exported to the market economy where they can do their work, the environmental economists using CV produces a machine that does work in the field producing an intermediate output (the WTP value) that can be exported to the planned economy. In this context, getting the appropriate WTP values naturally requires home-grown, not induced, preferences. Consequently, quantitative measurement plays a much more important role in CV.

Of course, sometimes experimental economists also want to measure a parameter like WTP, as in the case of the WTP for GMO-free foods discussed by Muniesa and Callon (2007). But such cases inevitably involve a good that can be delivered to participants. Consequently, the machine built to measure them does not need to be realistic. Rather than replicate a supermarket, the machine can be a highly artificial auction that has been shown to work well with induced values and then, so calibrated, can be turned loose in the lab to measure WTP. In that context, realism in the pursuit of an abstract parameter is no virtue. In contrast, environmental economists using CV typically cannot actually deliver the environmental quality at stake. Instead, they create a fictional counter-factional world in which consumers can purchase (or sell) the environmental resource of interest. Since the respondents do not have the incentive of actually obtaining the

good, CV researchers must maintain credibility by keeping the scenario as realistic as possible. They must tell a story in which there is a credible mechanism for improving the natural environment and a credible payment vehicle through which respondents might pay for it. In Randall et al.'s (1974) case, for example, this meant describing regulation of the power plant, showing photographs, and hypothetically paying through taxes or power bills.

Nevertheless, a common feature of both approaches is that the institutions (or machines) must be tested to ensure they are performing as intended. If not, they must be recalibrated or redesigned. Whereas experimental economics has succeeded by demonstrating that its institutions do work in the created environment, numerous tests seemed to indicate, at least to some, that CV was not performing as intended, that it was not creating economic man. A decades-long debate ensued over whether these were the result of particular CV practices or whether CV was itself inherently flawed. Moreover, in some cases, experiments designed to compare behavioral patterns in the laboratory under controlled conditions to the patterns observed with CV found no differences. To some practitioners of CV, this result was reassuring, as it seemed to confirm the problem was not inherent in CV. To others, however, the results were much more troubling: they seemed to suggest that *homo economicus* did not exist or could not be created.

At this point, anomalies in the CV literature began to attract the attention of behavioral economists.[8] Ironically, it was Jack Knetsch, in the 1960s and early '70s a leading figure in bringing the new welfare economics to the environmental literature, who, more than any other person, built a bridge from CV to behavioral economics. At some point in the late 1970s,

[8] For a history of behavioral economics, see Heukelom (2012) as well as Edwards' (2010) work on happiness economics. Surprisingly, the connection between behavioral economics and CV has been neglected in that literature, but I would suggest it was quite important, especially in the early 1980s as behavioral economics was gaining traction. A recent search on Google Scholar turned up 35 articles written by Dan Kahneman alone containing the term "contingent valuation." The sheer volume is instructive.

Knetsch became suspicious of nonmarket valuation and its behavioral foundations.[9] His partnership with Daniel Kahneman began in the early-mid 1980s with work on contingent valuation (as reported in Kahneman 1986) and their later work with Richard Thaler on the endowment effect remained motivated by problems in the CV literature (Kahneman, Knetsch, and Thaler 1990, 1991, Kahneman and Knetsch 1992). At this point, the CV literature had crossed a line. As long as the troublesome results from CV experiments undermined only CV, it could be relegated to a subfield of environmental economics and ignored by the rest of a skeptical profession. Once those results were perceived by some as undermining the postulates of rational economic behavior, CV attracted more powerful opposition.

At the risk of oversimplification, the various tests of CV and its institutions can be grouped into four key categories. The first went right to the question of strategic behavior posited by Samuelson (1954). Would people reveal their WTP even in the context of a market institution that was not "incentive compatible" (i.e. that gave no incentive for households to reveal their true values)? Early experimental work by Peter Bohm[10] (1972) and V. Smith (1979) seemed to imply that respondents did not free ride as much as Samuelson had predicted, and researchers sympathetic to the CV agenda such as Cicchetti and K. Smith (1973) and Cummings, Brookshire, and Schulze (1986) interpreted this finding as support for CV. However, these results were hardly persuasive to skeptics. In particular, failure to free-ride in a market setting (even one lacking incentive compatibility) was hardly the same as answering truthfully or accurately to a hypothetical question.

The second category of testing involved sensitivity analyses of the elicitation format.

[9] Growing skepticism is apparent in Gordon and Knetsch (1979).

[10] See Dufwenberg and Harrison (2008) on Bohm's place in experimental economics.

How much difference did it make how the WTP question was asked, and if it made a difference, which format performed better? Davis (1963a) and Randall, Ives, and Eastman (1974) had used a "bidding game," in which respondents are first asked to accept or reject a bid, and the bid is adjusted until they (just) accept it. However, research showed that the starting point in this bidding process biased the results, as respondents "anchored" on the starting value, and similar results were found for other methods. The open-ended approach ("how much would you be willing to pay…?") seemed to be the most unrealistic, to invite strategic responses, and to impose the heaviest cognitive burdens on respondents, who had to find their own WTP rather than react to one. By the 1990s, the consensus opinion of CV researchers converged around the dichotomous-choice format, in which respondents answered yes or no to a single WTP offer, perhaps with a single follow up. The approach had the downside of providing the least information (yes/no to a single value) and thus posing econometric challenging requiring various rival economic interpretations of the error terms. However, it obviously imposed the weakest cognitive burdens, was a familiar market institution (the posted price), and was incentive compatible (Mitchell and Carson 1989).

A third pair of tests surrounded so-called part-whole bias (or embedding). Part-whole bias is the potential for respondents to "embed" a whole set of concerns unrelated to the specific environmental resource at stake. For example, a study of water quality at a particular Lake A might, psychological, trigger thoughts about Lakes B and C, leading the respondent to value all three lakes instead of just A. In work circulating by 1984 and preliminarily reported in a comment by Kahneman (1986), Kahneman and Knetsch had found that WTP to stock fish in either of two lakes was similar to the WTP to stock all the lakes in Ontario.

The fourth category surrounded differences between WTP and WTA, an issue raised earlier by Randall and others. As a social construct, WTP fits a setting in which people must buy a commodity. In contrast, WTA fits a setting where people already own a commodity but may sell it. As a matter of abstract theory, any with-vs-without comparison can be valued using either WTP or WTA. From a Hicksian perspective, WTP and WTA differ by an income effect (the income associated with the different utility levels obtained). From a Coasean perspective, they differ by the property right: whether one came to the hypothesized market transaction owning the property or not. Standard theory suggests the income effect of the property right should be small, especially for most of the commodities hypothetically bought or sold in CV markets. For example, the money one would be willing to exchange for a permit to hunt geese should hardly be affected by the wealth associated with that permit. Famously, Willig (1976) showed that the WTP and WTA for *price* changes should be small. Randall and Stoll (1980) extended his results to exogenous changes in *quantities* of things consumed outside of markets, like environmental quality, a result that seemingly suggested the differences should be small in CV applications as well.

And yet, numerous CV studies found very large WTP-WTA disparities, on the order of 4x to 16x (Hammack and Brown 1974, Bishop and Heberlein 1979, Rowe, d'Arge, and Brookshire 1980). CV researchers and others had diverse reactions to these findings. One view was that Randall and Stoll misinterpreted their extension of Willig's results from price changes to quantity/quality changes, and that, in fact, there was nothing in economic theory that suggested WTP-WTA differences should be small (Hanemann 1991). A second view was that respondents rejected as implausible the WTA frame, and so used their responses to express moral outrage rather than market behavior.

A third view interpreted these results as part of a pattern consistent with the "endowment effect" proposed by Thaler (1980), namely, the idea that people valued things more once it was theirs or part of the status quo. Kahneman and Tversky's (1979) prospect theory was one way to explain this finding. Kahneman and Tversky argued that real human beings valued gains and losses differently and tended to avoid losses, for reasons that had nothing to do with the income effects that were needed to explain the differences according to neoclassical theory. In a series of laboratory experiments, Knetsch and Sinden (1984), Knetsch (1989), and Kahneman, Knetch, and Thaler (1990) found that WTP bids and WTA offers differed even in the lab, with induced preferences for raffle tickets or homegrown values for small items like candy bars and coffee mugs. These tests seemed to confirm the presence of an endowment effect. Importantly, while they clearly were motivated by the contingent valuation literature, they had much broader implications for human behavior.

To assess these and other issues with CV, in 1984 the US Environmental Protection Agency (EPA) sponsored a meeting organized by Ronald Cummings, David Brookshire, and William Schulze (Cummings et al. 1986). Cummings et al. provided a large volume of background material, and a group of expert environmental economists offered their perspectives on the matter, including Alan Randall, Richard Bishop and Thomas Heberlein, Myrick Freeman, and Kerry Smith. Of these, only Randall felt that CV was a method successfully proven to measure WTP, but the others were cautiously optimistic. A further group represented the larger economics profession, comprising Kenneth Arrow, Daniel Kahneman, Sherwin Rosen, and Vernon Smith. Arrow was the most "sympathetic" to the CV method (185). He viewed potential errors in the estimates of +/- 50%, such as had been documented by Cummings et al., as informative and within the same order of magnitude as errors in the natural sciences' ability to estimate the

physical effects of pollution.

Kahneman's response was nuanced and diplomatic. He was surprised the CV method performed as well as it did and accepted the exercise in principle. However, he urged more caution in thinking about the appropriate social setting in which the surveys were used. In particular, practitioners should think more about "the purchase structure" and whether a scenario involved, in people's minds, compensation for a loss or purchase of a gain (or avoidance of a loss). This distinction involved not only the property right inherent in the WTP/WTA comparisons, but also the cause of the changes in environmental quality to be evaluated and the reference point they would create for prospect theory. For example, he suggested that people would require compensation for the loss of a scenic view if caused by the government, but expect to pay to avoid its occurrence if caused by an insect. Kahneman felt that the data suggested that CV produced nonsensical results when constructing WTA, so should be restricted to WTP, but that WTP questions should not be imposed on compensation structures. In this way, he supported the use of surveys and CV for policy making, but in such a way that one expected neither to meet nor to construct economic man.

Like Kahneman, V. Smith did not expect CV to construct *homo economicus,* but for that reason he was much less enthusiastic about it. He cited an experiment in which individuals had bought and sold lottery tickets, providing bids and offers that could be interpreted as WTP and WTA. He showed that individuals' offered WTP and WTA were, in his words, "all over the map," even when incentivized, but that the market institution cleared at prices close to the inherent value. Drawing on that lesson, he concluded, "As I read the CVM work, it seems to me that what you are mainly working on is proposals for some sort of a substitution for the market, a calculation substitution. I really think we ought to devote a little time to thinking about whether

there might be the possibility of creating markets where they don't now exist, and let the market do the calculation" (202-3). He concluded with a proposal in which rights to resources could be assigned and people could bid for them. In other words, as with Plott's work on allocating airport slots (Svorenčík, this volume) Smith favored the design of real markets rather than using hypothetical markets to inform government planning.

In short, the perspectives of the wider profession at the 1984 CV conference reflected a reasonable cross-section of views: Arrow wanted to muddle through with benefit-cost analysis to inform policy as much as possible; V. Smith wanted to forego the entire CV project of "calculation substitution" for benefit-cost analysis and introduce market mechanisms into the real world; and Kahneman viewed CV through the lens of his brand of behavioral psychology rather than economics. The debate and discussion might have gone on in the same spirit, mirroring differences in the larger profession, while CV practitioners continued their applied work out of the spotlight. But meanwhile, economic abstractions like existence value entered benefic-cost and liability rules, upping the stakes for the applications.

IV. The *Exxon Valdez*

The increased stakes for CV developed through a series of US regulatory moves and court cases over, roughly, the decade of the 1980s. First, in 1979, the US Water Resources Council (supervising benefit-cost analyses of water projects) published its revised "Principles and Standards" for water resources planning, which allowed CV as one possible method for estimating recreation values. This precedent took on added importance when, in 1981, President Reagan issued Executive Order 12291, greatly expanding the role of benefit-cost analysis in agencies such as the EPA (K. Smith 1984). As with the earlier episodes for outdoor recreation, this gave agencies

like EPA added incentive to quantify ever more abstract and "intangible" benefits, such as the existence values identified by Krutilla.

That incentive was further extended to the Dept. of the Interior (DOI) by regulatory actions surrounding the 1980 Comprehensive Environmental Response Cleanup and Liability Act (CERCLA), better known as "Superfund." Superfund created a liability provision, known as the Natural Resource Damage Assessment (NRDA), whereby DOI appointed trustees who would sue potentially responsible parties for damages suffered by the public from hazardous releases. As DOI promulgated its NRDA rules, a key question centered around what constituted "damages." Importantly, Interior argued that the economic concept of consumer surplus was the appropriate conceptual framework for defining damages, but took a somewhat ambivalent position on the role of existence values. In 1989, a US Court of Appeals ruled in *State of Ohio v. US Dept. of Interior* that NRDAs should include existence values.

As noted by Cummings et al. (1986), these regulatory moves made CV a very valuable tool to federal agencies. Pushed to consider existence values, the agencies turned to CV as the only game in town for measuring such values.[11] However, a countervailing force soon came into play.

Just before the *Ohio* ruling came down, the *Exxon Valdez* ran aground in Alaska, spilling at least 11 million gallons of oil, the largest spill in US waters until the 2010 *Deepwater Horizon* spill. As the spill affected an area little used for recreation or other public activities but widely characterized as "pristine," existence values were, in the wake of *Ohio*, bound to play a large role

[11] Whereas Berman (this volume) suggests *de*regulation strengthened economists' position in US antitrust policy, I am suggesting the opposite is true in environmental economics. The difference may be that, in the environmental context, EPA's constituency valued the consumer benefits that come from regulation, whereas in the antitrust context they come from deregulation, but in either case documenting consumer benefits required economic analysis. An alternative view is that regulations beget deregulatory watchdogs (as with Reagan's executive order) and economics emerges in the synthesis.

in the case.[12] Indeed, the trustees sponsored a CV study of nonuse damages, which were estimated to be as high as $2.8 billion.[13] Rather than adopt the traditional NRDA strategy of critiquing the trustees' estimates of damages and supplying its own from its own economic experts, Exxon took a more radical approach: it invested large sums to fund research investigating the inherent flaws in CV as a measurement tool, and thus the impossibility of adequately supporting claims for nonuse values in court. (See Maas and Svorenčík 2016 for a discussion of this episode.)

Though Exxon settled out of court in 1991, its strategy already had turned to a second purpose when, in 1990, Congress passed the Oil Pollution Act, giving the National Oceanic and Atmospheric Administration (NOAA) a similar NRDA role for coastal oil spills. In promulgating its regulations, NOAA convened a blue ribbon panel to judge the academic research into CV and make recommendations about the role it should play in NOAA's NRDA regulations. Thus, Exxon's research, if released quickly, could help not only its own case but also influence the new NRDA regulations.

The research supported by Exxon was presented at a meeting held in 1992 and published, along with transcripts of the discussion, in Hausman (1993). Authors included, among others, Peter Diamond and Jerry Hausman, Daniel McFadden, environmental economists William Desvousges, Reed Johnson, and Kevin Boyle, and psychologists David Schkade and John Payne.[14] Commentators included Alan Randall, Zvi Griliches, Ken Arrow, and Charles Plott.

---

[12] As discussed by Fourcade (2011), these recent legal developments differentiated liability for oil spills in the US from France and countries, but were consistent with the older culture of US benefit-cost practices.

[13] The work was not published until much later (Carson et al. 2003).

[14] In interest of disclosure, I was working at Research Triangle Institute with Desvousges, Johnson, and Dunford during these years.

Numerous other economists were in attendance at the meetings, including Richard Bishop, Richard Carson, Michael Hanemann, Robert Mitchell, and Kerry Smith. Yet, discussion time was limited and tightly controlled. Additionally, many of the environmental economists present were offended by the claim made by some new to the issues that they were the first to introduce certain economic innovations, innovations that some of the environmental economists felt they had introduced long ago (see, e.g., the discussion in Hausman 1993 364-8). In general, the meeting became infamous as an occasion for bad feelings. As Portney (1994 3), put it, the entire post-Exxon debate was "spirited (and occasionally mean-spirited)."

Three particularly prominent studies introduced evidence that, in the authors' minds, proved CV to be fatally flawed. First, McFadden and Leonard (1993) found that different market institutions and question formats gave different WTP estimates, suggesting that current practices were not converging on the same construct. They also found that people seemed to anchor on the WTP suggested to them in dichotomous choice questions. Second, Diamond et al. (1993) applied CV to the value of preserving wilderness areas in Montana, using different versions with different combinations of areas threatened. To simplify, they theorized that, except for a small income effect, the WTP to preserve Area A when Areas A and B are both threatened plus the WTP to preserve Area B when only B is threated should equal the WTP to preserve Areas A and B. They found that CV results failed to meet this adding up test.

Third, Desvousges et al. (1993) found that CV studies were not sensitive to what they called "scope." They used CV to estimate the public's WTP to avoid deaths to migratory waterfowl from landing in uncovered oil ponds. Three versions measured damages from, respectively, the death of 200,000 birds, 20,000 birds, and 2,000 birds, but the WTP estimates were not significantly different across the three versions, a result similar to Kahneman and Knetsch's earlier

work on Ontario lakes. Following up on this work, Schkade and Payne (1993), the psychologists in the group, used the Desvousges et al. survey to conduct "verbal protocols," in which people think out loud during their responses to surveys. In response to the open-ended question in which they have to state their own maximum WTP, respondents indicated a great deal of confusion and difficulty arriving at a number, groping for various external cues to guide their thought. To them and other critics of CV, the worked implied that people do not have well-defined preferences that can be searched out and measured; rather, those preferences need to be constructed when encountering new decisions and decision contexts. That itself seemed a fatal flaw.

Respondents (e.g. Randall 1993, Hanemann 1994) argued that CV studies were reliable if done properly; if Exxon studies obtained inconsistent results it was because they were done improperly. For example, in some cases, Exxon researchers used open-ended question formats that were cognitively burdensome and were not incentive compatible, so the results of Schkade and Payne were hardly surprising. Additionally, they argued that many of the results for adding up tests, embedding, and WTP/WTA did not necessarily violate neoclassical preferences (Hanemann 1991, 1994, K. Smith 1992).

V. Conclusions

Backhouse and Biddle (2000) and Backhouse and Cherrier (2016), argue that the term "applied economics" is as ambiguous as it is ubiquitous. It has meant the application of abstract theoretical models to problems grounded in the real world, has been a simple synonym for empirical work, and, I would add, sometimes has been a euphemism for mediocrity. As they note, the ambiguity arises because, ultimately, "applied economics" has been defined in opposition to "pure," "theoretical," or "abstract" economics. A corollary to this observation is that the relationship be-

tween applied economics and the broader field of economics is also ambiguous. The varied reactions to the CV debate serve to illustrate the range of potential interpretations for this relationship when applied work does not go as expected. Consider three key positions.

1. Interpret CV as a successful application of economic theory.

Economists adopting this perspective have argued one can draw lessons from the observed anomalies, to learn which CV practices construct WTP values that look rational from the standpoint of economic theory. Michael Hanemann (1994), for example, was willing to concede that people do not have pre-existing preference mappings that can simply be accessed by a survey. To him, "the real issue is not whether preferences are a construct but whether they are a *stable* construct" (28). After all, people also construct their preferences in new market settings. Here, Hanemann quotes Robert Solow's reaction to the Schkade-Paine verbal protocols, to wit, they "sound a lot like Bob Solow in the grocery store". Thus, from this perspective, hypothetical CV markets, like other markets, have the potential to construct meaningful values. Here, the applied work is a (potentially) successful application of economic theory.

When it issued its report, NOAA's blue ribbon panel (Kenneth Arrow; Robert Solow; Paul Portney, a leading environmental economist; Edward Leamer; Roy Radner; and Howard Schuman, a survey researcher) took a similar line (Arrow et al. 1993). After weighing the evidence, the panel concluded that "CV studies convey useful information. We think it is fair to describe such information as reliable by the standards that seem to be implicit in similar contexts, like market analysis for new and innovative products and the assessment of other damages normally allowed in court proceedings" (4610). Note here the somewhat faint praise. Consistent with Arrow's "sympathy" at the 1984 EPA conference, CV in this view is hardly a clean application of economic theory; it is messy, but good enough for government work. Additionally, they

offered a number of specific suggestions, all intended to ensure the values of *homo economicus* were constructed.

2. Conclude that the anomalies found in CV studies are evidence of problems with economics.

A second reaction is at the opposite end of the spectrum from the first, rejecting the existence of economic man, at least in some important contexts. As we have already seen, Jack Knetsch began to travel down this round sometime around 1980. More recently, Daniel McFadden, who won a Nobel Prize in economics for his work uniting theories of utility-maximization to the econometric analysis of choices, has all but abandoned the economic theory of the utility-maximizing consumer (2009, 2013). McFadden's own journey down this road clearly was underway at the Exxon conference. He commented in the discussion that "if people really have this willingness-to-pay, [the values] should shine through" all the variations in the CV studies (Hausman 1993 214). McFadden clearly was troubled by the Exxon results. In 1999, he published an article titled "Rationality for Economists?" in which he argued that *homo economicus*, or what he called "Chicago man," "is an endangered species." Notably, he suggested that public goods were a case in point and that "nowhere has this been more evident than in economists' attempts to value non-use public goods, such as endangered species or wilderness areas" (97). He concluded by arguing that, for economics, "the challenge is to evolve Chicago man in the direction of K-T [Kahneman-Tversky] man" (99).

Here, the relationship between the applied economics and the economics generally is ambiguous. On one level, applications and theory are in dialogue, perceived failures to apply one theory lead to a new one. On another level, there is already a pre-existing profound disagreement here between behavior and neoclassical economists on what a good economic theory looks like, and which might be being applied is itself a contested question.

3. Interpret CV as a failed attempt to apply economic theory, or even distance it from economics.

Reacting to this more radical critique, a third interpretation would protect the integrity of neoclassical economics by severing from it the difficulties raised by CV. For example, in his comments at the Exxon conference, Plott (1993) argued that the evidence of WTP/WTA disparities and embedding were not "simply a measurement problem" (468). "The problem here is that the fundamental consistency property, the transitivity of preferences that is required of all theories of optimization, is absent" (471). He concluded that CV does not "measure what the underlying philosophy of welfare economics and preference theory requires it to measure" (468). Potentially going further, he raised the possibility that those theories are not "sufficiently well-grounded in scientific evidence" (468), but pulled back from that conclusion. Evidence pointing that way should be ignored, for now, he said, on the grounds that welfare theory is too far rooted in real-world institutions to uproot until there is a better theory to replace it. But he argued that "economists must begin to deal systematically with this … issue" (477).

Plott has taken his own advice and pursued just such a research agenda in the ensuing years. Leaving the CV debate aside except for a few passing comments, Plott has turned his attention to the more threating evidence from the Kahneman-Knetsch line of studies showing similar WTP/WTA disparities and other anomalies in the laboratory with real payments. In a series of studies, he has argued that WTP/WTA disparities in the environmental literature and elsewhere are not robust to experimental procedures but that they disappear with incentive-compatible mechanisms and when subjects are given practice and training (Plott and Zeiler 2005), that Knetsch's finding on preference reversals can be explained by incorporating preferences for gifts rather than loss aversion and prospect theory (Plott and Zeiler 2007), and that anomalies may lie

with failure to understand the market institution (Cason and Plott 2014). In other words, where CV failed to create economic man, other institutions still could.

Peter Diamond and Jerry Hausman went further. They concluded their 1994 *Journal of Economic Perspectives* article by arguing that

> contingent valuation is a deeply flawed methodology for measuring nonuse values, one that does not estimate what its proponents claim to be estimating. It is precisely in the lack of experience both in markets for environmental commodities and in the consequences of such decision that makes contingent valuation questions so hard to answer and the responses so suspect. (62)

In other words, the hypothetical nature of the market, which created the need for CV in the first place, precludes the possibility of measuring values. Still speaking in the context of the CV debate, Hausman, never one to mince words, later suggested that "environmental economics is to economics what military music is to music" (Coy 1997)—thus separating the applied from the pure.

The history of CV is the history of an effort to produce *homo economicus* where no markets exist. As with experimental economics (an intellectual cousin), this effort entailed creating an economy (here, a hypothetical one), in which people could act and in which economists could observe their actions and measure their WTP values. As CV became further embedded in US policy-making processes, the pressure increased to get CV right, where "right" meant consistent with a microeconomic theory of consumer optimization. However, CV produced a number of anomalies that seemingly could not be reconciled with such a theory. To some, it appeared that, once turned loose in these hypothetical economies, people did not behave like *homo economicus*.

Though the CV debate remains far from settled (Haab et al. 2013), it has produced several possible positions, from continuing the effort to tweak CV to rejecting the theory of economic man. The varied reactions to the CV debate illustrate the range of potential interpretations for the relationship of applied economics to economics when applied work does not go as expected.

References

Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman. 1993. NOAA Panel on Contingent Valuation. 15 CFR Chapter IX, Federal Register 58: 4601-14.

Backhouse, Roger E., and Jeff Biddle. 2000. "The Concept of Applied Economics: a History of Ambiguity and Multiple Meanings," *HOPE* 32 (annual supplement): 1-24.

Backhouse, Roger E., and Béatrice Cherrier. 2016. "The Age of the Applied Economist: The Transformation of Economics since the 1970s." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2868144

Banzhaf, H. Spencer. 2009. Objective or Multiobjective? Two Historically Competing Visions for Benefit Cost Analysis. *Land Economics* 85: 1-23.

Banzhaf, H. Spencer. 2010. Consumer Surplus with Apology. *Annual Review of Resource Economics* 2: 183-207.

Banzhaf, H. Spencer. 2014. Retrospectives: The Cold-War Origins of the Value of Statistical Life. *Journal of Economic Perspectives* 28(4): 213-26.

Banzhaf, H. Spencer. 2016. The Environmental Turn in Natural Resource Economics: John Krutilla and "Conservation Reconsidered." Mimeo.

Bishop, Richard C., and Thomas Heberlein. 1979. Measuring Values of Extra-Market Goods: Are Indirect Measures Biased? *American Journal of Agricultural Economics* 61(5): 926-30.

Bohm, Peter. 1972. "Estimating Demand for Public Goods: An Experiment." *European Economic Review* 3(2): 111-130.

Boulier, Bryan L., and Robert S. Goldfarb. 1998. On the Use and Nonuse of Surveys in Economics. *Journal of Economic Methodology* 5(1):1-21.

Boumans, Marcel J. 2005. *How Economists Model the World into Numbers*. New York: Routledge.

Bowley, Arthur L. 1920. Cost of Living and Wage Determination. *Economic Journal* 30(117): 114-7.

Brookshire, David S., Alan Randall, and John R. Stoll. 1980. Valuing Increments and Decrements in Natural Resource Service Flows. *American Journal of Agricultural Economics* 62(3): 478-88.

Callon, Michel, ed. 1998. *The Laws of Markets*. Oxford: Blackwell.

Carson, Richard T. 2011. *Contingent Valuation: A Comprehensive Bibliography and History*.

Cheltenham, UK: Edward Elgar.

Carson, Richard T., Robert C. Mitchell, Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud. 2003. Contingent valuation and lost passive use: damages from the Exxon Valdez oil spill. *Environmental and Resource Economics* 25(3): 257-86.

Cason, Timothy N., and Charles R. Plott. 2014. Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing. *Journal of Political Economy* 122(6): 1235-1270.

Cicchetti, Charles J., and V. Kerry Smith. 1973. Congestion, Quality Deterioration, and Optimal Use: Wilderness Recreation in the Spanish Peaks Primitive Area. *Social Science Research* 2: 15-30.

Ciriacy-Wantrup, S.V. 1947. Capital Returns from Soil Conservation Practices. *Journal of Farm Economics* 29: 1181-96.

Coy, Peter. 1997. "The In Your Face Economist." *Businessweek*. Accessed at http://www.bloomberg.com/news/articles/1997-06-29/the-in-your-face-economist.

Cummings, Ronald G., David S. Brookshire, and William D. Schulze, eds. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, NJ: Rowman & Allanheld.

Davis, Robert K. 1963a. *The Value of Outdoor Recreation: An Economic Study of the Maine Woods*. Dissertation, Harvard University.

Davis, Robert K. 1963b. Recreation Planning as an Economic Problem. *Natural Resources Journal* 3: 239-49.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara P. Hudson, and K. Nicole Wilson. 1993. *Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Responsibility*. In Hausman (1993): 91-164.

Diamond, Peter A., and Jerry A. Hausman. 1994. Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8(4): 45-64.

Diamond, Peter A., Jerry A. Hausman, Gregory K. Leonard, and Mike A. Denning. 1993. Does Contingent Valuation Measure Preferences? Experimental Evidence. In Hausman (1993): 41-90.

Dufwenberg, Martin, and Glenn W. Harrison. 2008. Peter Bohm: Father of Field Experiments. *Experimental Economics* 11(3): 213-220.

Edwards, José M. 2010. *Joyful Economists: Remarks on the History of Economics and Psychology from the Happiness Studies Perspective*. PhD Dissertation, Université Paris I Panthéon-Sorbonne.

Fourcade, Marion. 2011. Cents and Sensibility: Economic Valuation and the Nature of "Nature." *American Journal of Sociology* 116(6): 1721-77.

Gordon, Irene M., and Jack L. Knetsch. 1979. Consumer's Surplus Measures and the Evaluation of Resources. *Land Economics* 55(1): 1-10.

Guala, Franceso. 2005. *The Methodology of Experimental Economics*. New York: Cambridge University Press.

Guala, Franceso. 2007. How to Do things with Experimental Economics. In *Do Economists Make Markets? On the Performativity of Economics*, ed. by Donald MacKenzie, Fabian Muniesa, and Lucia Siu. Princeton: Princeton University Press, 128-62.

Haab, Timothy C., Matthew G. Interis, Danield Petrolia, and John C. Whitehead. 2013. From Hopeless to Curious? Thoughts on Hausman's "Dubious to Hopeless" Critique of Contingent Valuation. *Applied Economic Perspectives and Policy* 35(4): 593-612.

Hammack, Judd, and Gardner M. Brown. 1974. *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Baltimore: Johns Hopkins U. Press for Resources for the Future.

Hanemann, W. Michael. 1991. Willingness-to-Pay vs. Willingness-to-Accept: How Much Can They Differ? *American Economic Review* 81(3): 635-47.

Hanemann, W. Michael. 1992. Preface. In *Pricing the European Environment*, ed. by Ståle Navrud. New York: Oxford University Press, 9-35.

Hanemann, W. Michael. 1994. Valuing the Environment through Contingent Valuation. *Journal of Economic Perspectives* 8(4): 19-43.

Hausman, Jerry A., ed. 1993. *Contingent Valuation: A Critical Assessment*. Amsterdam: North Holland.

Heukelom, Floris. 2012. Three explanations for the Kahneman-Tversky Programme of the 1970s. *The European Journal of the History of Economic Thought* 19(5): 797-828.

Kahneman, Daniel. 1986. Comments by Professor Daniel Kahneman. In Cummings, Brookshire, and Schulze (1986), 185-94.

Kahneman, Daniel, and Jack Knetsch. 1992. Valuing Public Goods: The Purchase of Moral Satisfaction. *Journal of Environmental Economics and Management* 22: 57-70.

Kahneman, Daniel, Jack Knetsch, and Richard H. Thaler. 1990. Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy* 98(6): 1325-48.

Kahneman, Daniel, Jack Knetsch, and Richard H. Thaler. 1991. Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives* 5(1): 193-206.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2): 263-291.

Knetsch, Jack L. 1989. The Endowment Effect and Evidence of Nonreversible Indifference Curves. *American Economic Review* 79(5): 1277-84.

Knetsch, Jack L., and Robert K. Davis. 1966. Comparison of Methods for Recreation Evaluation. In *Water Research*, ed. by Allan V. Kneese and Stephen C. Smith. Washington: John Hopkins Press for Resources for the Future, 125-42.

Knetsch, Jack L., and J.A. Sinden. 1984. Willingness to Pay and Compensation Demand: Experimental Evidence of an Unexpected Disparity in Measures of Value. *Quarterly Journal of Economics* 99(3): 507-21.

Krutilla, John V. 1967. "Conservation Reconsidered." *American Economic Review* 57(4): 777-86.

Maas, Harro, and Mary S. Morgan, eds. 2012. *Observing the Economy*. Durham, NC: Duke U. Press.

Maas, Harro, and Andrej Svorenčík. 2016. Fraught with Controversy: Organizing Expertise on Contingent Valuation. Mimeo.

MacKenzie, Donald. 2007. Is Economics Performative? Option Theory and the Construction of Derivatives Markets. In *Do Economists Make Markets? On the Performativity of Economics*, ed. by Donald MacKenzie, Fabian Muniesa, and Lucia Siu. Princeton: Princeton University Press, 54-86.

McCloskey, Donald. 1983. The Rhetoric of Economics. *Journal of Economic Literature* 21(2): 481-517.

McFadden, Daniel. 1999. Rationality for economists? *Journal of risk and uncertainty* 19: 73-105.

McFadden, Daniel. 2013. The *New* Science of Pleasure: Consumer Choice Behavior and the Measurement of Well-Being. In *Handbook of Choice Modelling,* ed. by S. Hess and A.J. Daly. Edward Elgar.

McFadden, Daniel, and Gregory K. Leonard. 1993. Issues in the Contingent Valuation of Environmental Goods. In Hausman (1993): 165-216.

Medema, Steven G. 2014. The Curious Treatment of the Coase Theorem in the Environmental Economics Literature, 1960–1979. *Review of Environmental Economics and Policy* 8(1): 39-57.

Mitchell, Robert Cameron, and Richard T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.

Moscati, Ivan. 2007. Early Experiments in Consumer Demand Theory: 1930-1970. *HOPE* 39: 359-401.

Muniesa, Fabian, and Michel Callon. 2007. Economic Experiments and the Construction of Markets. In *Do Economists Make Markets? On the Performativity of Economics*, ed. by Donald MacKenzie, Fabian Muniesa, and Lucia Siu. Princeton: Princeton University Press, 163-89.

Plott, Charles R. 1993. Contingent Valuation: A View of the Conference and Associated Research. In Hausman (1993): 467-78.

Plott, Charles R., and Kathryn Zeiler. 2005. The willingness to pay-willingness to accept gap, the 'endowment effect,' subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review* 95(3): 530-45.

Plott, Charles R., and Kathryn Zeiler. 2007. Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review* 97(4): 15-38.

Portney, Paul R. 1994. The Contingent Valuation Debate: Why Economists should Care. *Journal of Economic Perspectives* 8(4): 3-17.

Raiffa, Howard. 1969. Preferences for Multi-Attributed Alternatives. Memorandum RM-5868-DOT/RC. Prepared for U.S. Department of Transportation, Federal Railroad Administration, Office of High Speed Ground Transportation. The RAND Corporation, Santa Monica, California.

Randall, Alan. 1972. Market Solutions to Externality Problems: Theory and Practice. *American Journal of Agricultural Economics* 54(2): 175-83.

Randall, Alan, Berry Ives, and Clyde Eastman. 1974. Bidding Games for Valuation of Aesthetic Environmental Improvements. *Journal of Environmental Economics and Management* 1: 132-49.

Randall, Alan, and John R. Stoll. 1980. Consumer's Surplus in Commodity Space. *American Economic Review* 70(3): 449-55.

Rowe, Robert D., Ralph C. d'Arge, and David S. Brookshire. 1980. An Experiment on the Economic Value of Visibility. *Journal of Environmental Economics and Management* 7: 1-19.

Samuelson, Paul A. 1954. The Pure Theory of Public Expenditures. *Review of Economics and Statistics* 36(4): 387-9.

Schelling, T.C. 1968. "The Life You Save May be Your Own." In Problems in Public Expenditure Analysis, ed. by Samuel B. Chase, Jr. Washington, DC: Brookings Institution.Schkade, David A., and John W. Payne. 1994. How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental

regulation. *Journal of Environmental Economics and Management* 26(1): 88-109.

Smith, V. Kerry, ed. 1984. *Environmental Policy under Reagan's Executive Order: The Role of Benefit-Cost Analysis.* Chapel Hill: University of North Carolina Press.

Smith, V. Kerry. 1992. Arbitrary Values, Good Causes, and Premature Verdicts. *Journal of Environmental Economics and Management* 22(1): 71-89.

Smith, Vernon L. 1979. Incentive Compatible Experimental Processes for the Provision of Public Goods. In *Research in Experimental Economics*, ed. by Vernon L. Smith. JAI Press.

Smith, Vernon L. 1982. Markets as Economizers of Information: Experimental Examination of the "Hayek Hypothesis." *Economic Inquiry* 20: 165-79.

Stigler, George J. 1943. The New Welfare Economics. *American Economic Review* 33(2): 355-9.

Svorenčík, Andrej. 2015. *The Experimental Turn in Economics: A History of Experimental Economics*. Dissertation, University of Utrecht.

Thaler, Richard. 1980. "Toward a Positive Theory of Consumer Choice." Journal of Economic Behavior & Organization 1(1): 39-60.

Willig, Robert D. 1976. Consumer's Surplus without Apology. *American Economic Review* 66(4): 589-597.