

The Stated Preference Method

IS IT RELIABLE FOR DETERMINING NATURAL
RESOURCE DAMAGES?

WHAT ARE THE ALTERNATIVES?

PRESENTED BY

Mark Berkman
David Sunding
Dan McFadden
Vic Adamowicz
Loren Dunn
James Bruen

September 27, 2018

THE **Brattle** GROUP

Agenda

- **Overview of the Issue and the CV Method**
- Review of the State of the Method
- Unmet Concerns Regarding the Method
- Review of the Alternative Tools
- Legal Perspectives
- Panel Discussion and Q&A

Agenda

- Overview of the Issue and the CV Method
- **Review of the State of the Method**
- Unmet Concerns Regarding the Method
- Review of the Alternative Tools
- Legal Perspectives
- Panel Discussion and Q&A

THE STATED PREFERENCE METHOD: CONTEMPORARY GUIDANCE AND NEXT STEPS FOR RESEARCH AND PRACTICE

Vic Adamowicz
University of Alberta, Canada

Johnston, R., K. Boyle, W. Adamowicz, J Bennett, R. Brouwer, T.A. Cameron, W.M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, C. Vossler. 2017. Contemporary Guidance for Stated Preference Studies. **Journal of the Association of Environmental and Resource Economists**. 4(2): 319-405.

Thanks to Rob Johnston, Kevin Boyle and Pat Lloyd-Smith for comments on this presentation.

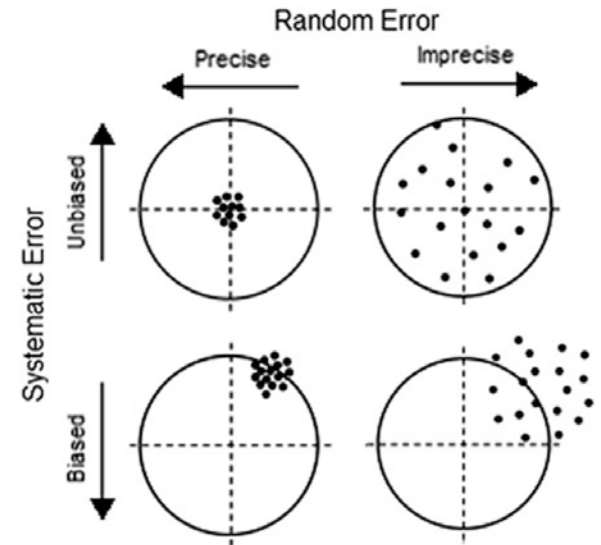
Definitions

- Value: Amount of money (paid or received) that makes one as well off as before a change (in quality, quantity, price, etc.)
- Public Goods: Everyone can benefit from these, consumption by one person doesn't reduce anyone else's consumption (non-rival, non-excludable)
- Private Goods: Rival, excludable
- Use Values: Values that can be measured from observable behavior
- Passive Use Values: Values that do not have an observable behavioral trail
- Stated Preference: can measure use and passive use values
- Revealed Preference: can measure only use values

Validity

- Willingness to pay (WTP), or value, is *unobservable*
- We can observe choices, at different prices, qualities, etc., and by making various assumptions, derive measures of WTP.
- We can ask structured questions, and making various assumptions, derive measures of WTP.
- How do we assess validity?
- Bishop and Boyle (2017; 2018)

Fig. 12.1 Reliability and validity illustrated. *Source* https://www.ucl.ac.uk/ich/short-courses-events/about-stats-courses/stats-rm/Chapter_9_Content/relation_between_reliability_validity



Validity

- Reliability versus validity
- Content validity: design / implementation
- Construct validity: responds as expected in theory
- Convergent validity: similar to measures with the same theoretical underpinnings designed to measure the same value
- Criterion validity: related to measures of “criteria” / “truth”
- Evidence that stated preference methods *can* produce valid results – but under what conditions?
- Bishop and Boyle (2017; 2018)

Elements of Validity

- Survey development (pre-testing, defining the good, provision mechanism, payment vehicle, etc.)
- Value elicitation (incentive compatibility, consequentiality, randomization of prices / attributes)
- Data collection and analysis
- Validity assessment
- Reporting
- Evaluation using the weight of evidence

Evolution of the Literature – The past decade

- Value Elicitation: mechanism design, consequentiality, and incentive compatibility
- Survey modes
- Presentation of valuation tasks (attributes, alternatives)
- Empirical / econometric methods

Evolution of the Literature: Value Elicitation

- **There has been a “sea change” in the Stated Preference literature in the past decade (Johnston et al, 2017)**
 - Survey participants respond to incentives / perceive consequences.
 - Expectations about provision of the good
 - Expectations about the payment being requested
 - Differences between private goods and public goods
 - Under certain conditions, a stated preference question is “incentive compatible” (dominant strategy of the individual is to provide true WTP)

Conditions for Incentive Compatible Elicitation for Public Goods: Vossler et al (2012), pg 151

Summary of Carson and Groves (2007):

- i. the participants care about the outcome;
- ii. the authority can enforce payments by voters;
- iii. the elicitation involves a yes or no vote on a single project; and
- iv. the probability that the proposed project is implemented is weakly monotonically increasing with the proportion of yes votes.

Evolution of the Literature:

- Stated and revealed preference methods can suffer from various challenges: incentive problems, measurement error, etc.
- How does one “test” stated preference methods?
 - Not as simple as comparing stated preference to a purchase in an experiment.....
- One “test” for public goods valuation – compare stated preferences to a referendum
- Or test using economic laboratory experiments

Is hypothetical bias universal? Validating contingent valuation responses using a binding public referendum ☆

Robert J. Johnston*

476

R.J. Johnston / Journal of Environmental Economics and Management 52 (2006) 469–481

Table 2

Raw results: proportion of 'yes' votes at different quarterly cost levels

	Proportion of 'Yes' votes		Test of equal proportions (H_0 : Survey proportion = Referendum proportion)
	Hypothetical survey responses	Binding referendum responses	
Quarterly cost level ^a			
\$125 ($n = 72$)	0.653	—	
\$175 ($n = 75$)	0.587	—	
\$250 ($n = 64$ survey; $n = 267$ referendum)	0.484	0.457	$Z = 0.396, p = 0.692$
\$325 ($n = 71$)	0.409	—	
\$425 ($n = 46$)	0.326	—	

^aNumbers in parentheses indicate the total number of surveys received at each stated cost level.

An Experimental Example: Carson et al, 2014

Table 2. Experimental Data

Regime	Treatment	Individual Vote Summary	
		% Yes	Total Subjects
Binding majority	Real mechanism 1 (100% chance of binding)	45.8	96
Binding majority	$P(80\%)$ mechanism 1 (80% chance of binding)	41.3	46
Binding majority	$P(50\%)$ mechanism 1 (50% chance of binding)	48.1	52
Binding majority	$P(20\%)$ mechanism 1 (20% chance of binding)	44.0	50
Binding majority	Inconsequential (0% chance of binding)	60.3	58
Binding majority with free chance	$PF(80\%)$ mechanism 2 (for passed proposition: 80% chance of payment, 20% chance of free provision)	58.2	55
Binding majority with free chance	$PF(20\%)$ mechanism 2 (for passed proposition: 20% chance of payment, 80% chance of free provision)	71.4	49

Validity

- Validity is a concern in any environmental valuation study – revealed preference or stated preference
- There are many components of a study that can improve validity, and many research questions remain
 - Survey development (pretesting, defining the good, provision mechanism, payment vehicle, etc.)
 - Value elicitation (consequentiality, incentive compatibility, randomization)
 - Data collection and analysis
 - Validity assessment
 - Reporting
- There is a substantial literature showing that stated preference methods can be valid, but validity depends on the study design and implementation

Research Areas

- Continued assessment of validity and reliability (of all valuation methods)
- Role of incentive compatibility and relevance across different WTP elicitation formats for public goods
- Incentive compatibility for private goods
- Exploration of Consequentiality
- WTP versus WTA
- Payment vehicles
- Individuals versus households
- Sampling, survey mode, response bias, attention, and related aspects
- Communication / description of the “good” being valued
- “How good is good enough”?

References:

- Bishop, R. and K. Boyle. 2017. Reliability and validity in nonmarket valuation. In *A Primer on Nonmarket Valuation*. Springer. Dordrecht.
- Bishop, R. and K. Boyle. 2018. Reliability and validity in nonmarket valuation. *Environmental and Resource Economics*.
- Carson, Richard T., and Theodore Groves. 2007. Incentive and informational properties of preference questions. *Environmental and Resource Economics* 37:181–210.
- Carson, Richard T., Theodore Groves, and John A. List. 2014. Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists* 1:171–207.
- Johnston, Robert J. 2006. Is hypothetical bias universal? Validating contingent valuation response using a binding public referendum. *Journal of Environmental Economics and Management* 52:469–81.
- Johnston, R., K. Boyle, W. Adamowicz, J Bennett, R. Brouwer, T.A. Cameron, W.M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, C. Vossler. 2017. Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists*. 4(2): 319-405.
- Lloyd-Smith, P. and W. Adamowicz Accepted. Can Stated Measures of Willingness-to-Accept be Valid? Evidence from Laboratory Experiments. *Journal of Environmental Economics and Management*.
- Parker, J A and N S Souleles. 2017. Reported preference v. revealed preference: Evidence from the propensity to spend tax rebates. NBER Working Paper w23920.
- Vossler, Christian A., Joe Kerkvliet, Stephen Polasky, and Olesya Gainutdinova. 2003. Externally validating contingent valuation: An open-space survey and referendum in Corvallis, Oregon. *Journal of Economic Behavior and Organization* 51:261–77.
- Vossler, Christian A., Maurice Doyon, and Daniel Rondeau. 2012. Truth in consequentiality: Theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics* 4:145–71.

Agenda

- Overview of the Issue and the CV Method
- Review of the State of the Method
- **Unmet Concerns Regarding the Method**
- Review of the Alternative Tools
- Legal Perspectives
- Panel Discussion and Q&A

The Stated Preference Landscape

1. Contingent Valuation Method (Referendum format): Subjects are asked if they would vote for a specified environmental policy (e.g., cleanup oiled beaches) at a specified cost to them

- Variants: Double referendum, Open-ended
- Issues: Anchoring, Warm Glow, Scope and Extension Neglect

2. Choice-Based Conjoint Analysis: Subjects are asked to make choices from a series of menus with various policies and policy costs

- Variants: Ranking or Rating Alternatives, Elicitation of Compensating Adjustments
- Issues: Reference Point Bias, Manipulation in Training/Context, Shadowing and Fatigue
- Pros: Opportunities for internal testing for consistency, calibration (to past market or social choices that are considered sound and comparable precedents)

SP Methods – The Reliability Gradient

MORE ← RELIABLE → LESS

Application	RP calibration possible?	Forecasting: Consumer choice in market or non-market decisions	Welfare analysis: Remediation of harm
Familiar market goods with purchase experience	Yes	Yes	Yes
Unfamiliar market goods without prior experience	Yes	Noisy	Noisy
Use of familiar services (e.g., fishing) provided by environmental resources, not marketed directly	Yes	Yes	Yes
Valuation of unfamiliar non-use services provided by environmental resources, no market	No	Internal consistency checks ⇒ unreliable	Inadequate foundation

↑ MORE
↑ RELIABLE
↓ LESS

CBC Checklist

	Issue
1	Familiarity: Subjects should be familiar with the class of products studied and their attributes, after training if necessary, and experienced in making market choices among them.
2	Sampling, Recruitment, Background: Subjects should be sampled randomly from the target population, and compensated sufficiently to assure participation and attention. Background on socioeconomic status and purchase history should be collected for sample correction and modelling of heterogeneity in choice behavior.
3	Outside Option: Menus should include a meaningfully specified “no purchase” alternative, or the study should be designed to use equivalently identifying external background or real market data.
4	Menu Design: The number of menus, products per menu, attributes per product, and span and realism of attribute levels must balance market realism and the independent variation needed for statistical accuracy.
5	Attribute formatting: The clarity and prominence of attributes should mimic, to the extent possible given the goals of the analysis, the information environment the consumer will face in the real market.
6	Elicitation Frame: Elicitation formats other than hypothetical market choice must balance information value against the risk of invoking incompatible cognitive frames.
7	Incentive Alignment: Elicitations with a positive incentive for a truthful response reduce the risk of carelessness or casual opinion.
8	Subject Training: Training may be necessary to provide sufficient familiarity and experience with products, but to minimize bias due to unrealistic conveyance of information, it needs to mimic the “street” training provided by the real market.
9	Calibration and Testing: Consistency and reality checks should be built into the study, and forecasting accuracy should, if possible, be tested against external data.

McFadden & Train (2017) *Contingent Valuation of Environmental Goods*

1. **Undersensitivity to costs** – Estimated WTP can vary by large factors with different cost prompts
2. **Anchoring to high-cost cues** – Estimated WTP rises with level of highest cost prompt; see also Green, Jacowitz, Kahneman, McFadden (1998)
3. **Extension neglect on payments** – Subjects focus on stated dollar cue, whether stated per month, per year, single or repeating payment
4. **Extension neglect and scope test on harm** – stated preferences appear driven primarily by concept of improvement rather than amount of improvement
5. **The meaning of “No”** – What are the future options available to a subject if they vote against the referendum posed in the CV study?

Are “Bad” CV Studies “Good” for the Environment?

1. NOAA study of reef protection (Bishop et al, 2011) estimates repairing 5 acres of reef provides a social benefit of \$7.3 billion, at a cost of at most \$13.2 million, all figures per year.
 - If this CV study is considered reliable, then a responsible party in a case of environmental harm can rightfully claim \$7.3 billion in compensatory restoration by spending \$13.2 million on reef repair.
2. The Deepwater Horizon oil spill was estimated in a CV study (Bishop et al, 2016) to have a WTP to avoid a future spill of \$15.3 billion to \$17.2 billion.
 - The responsible parties in the Deepwater Horizon spill could fully compensate the public by paying \$13.2 million for reef repair for 3 years, less than \$40 million total.

CONCLUSION: “Bad” CV (or CBC) studies lead to inconsistent environmental policy and incentive distortions that are bad for the environment.

Agenda

- Overview of the Issue and the CV Method
- Review of the State of the Method
- Unmet Concerns Regarding the Method
- **Review of the Alternative Tools**
- Legal Perspectives
- Panel Discussion and Q&A

Agenda

- Overview of the Issue and the CV Method
- Review of the State of the Method
- Unmet Concerns Regarding the Method
- Review of the Alternative Tools
- **Legal Perspectives**
- Panel Discussion and Q&A

Legal Views



James A. Bruen

Partner

Farella Braun + Martel LLP
San Francisco

jbruen@fbm.com



Loren Dunn

Principal

Beveridge & Diamond
Seattle

ldunn@bdlaw.com

Panel Discussion and Q&A

THE POWER OF **ECONOMICS**

brattle.com

THE **Brattle** GROUP

