



School of Business,
Economics and Law
UNIVERSITY OF GOTHENBURG

WORKING PAPERS IN ECONOMICS

No 418

Design of stated preference surveys: Is there more to learn from behavioral economics?

Fredrik Carlsson,

November 2009

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

SCHOOL OF BUSINESS, ECONOMICS AND LAW, UNIVERSITY OF GOTHENBURG

Department of Economics

Visiting adress Vasagatan 1,

Postal adress P.O.Box 640, SE 405 30 Göteborg, Sweden

Phone + 46 (0)31 786 0000

Design of stated preference surveys: Is there more to learn from behavioral economics?

Fredrik Carlsson^a

Abstract

We discuss the design of stated preference (SP) surveys in light of findings in behavioral economics such as context dependence of preferences, learning, and differences between revealed and normative preferences. More specifically, we discuss four different areas: (i) revealed and normative preferences, (ii) learning and constructed preferences, (iii) context dependence, and (iv) hypothetical bias. We argue that SP methods would benefit from adapting to some of the findings in behavioral economics, but also that behavioral economics may gain insights from studying SP methods.

Key words: stated preferences, behavioral economics.

JEL codes: C91, D03, H4, Q51

Acknowledgements: I wish to thank Francisco Alpizar, Olof Johansson-Stenman, Mitesh Kataria, Elina Lampi, Peter Martinsson and participants at the conference Behavioral economics: What can it contribute to environmental and resource economics? at Western Washington University for comments and discussions..

^a Department of Economics, School of Business, Economics and Law, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden; Ph +46 31 773 41 74; E-mail fredrik.carlsson@economics.gu.se Department of Economics, University of Gothenburg, Box 640, SE 405 30 Gothenburg, Sweden.

1. Introduction

The field of behavioral economics has grown rapidly in the last ten years. The background is a wealth of evidence, often experimental, identifying empirical phenomena that are not adequately explained by traditional economic analysis. Behavioral economics explores these anomalies and develops models that incorporate factors such as emotions, fairness, reciprocity, social norms, and bounded rationality. The stated preference literature was early on influenced by both behavioral and experimental economics; and perhaps behavioral economics to some extent has been influenced by the stated preference literature.¹ One reason for the interest in behavioral economics was most likely the anomalies found in applied work including stated preference studies. For example, the empirical findings regarding the huge differences between WTP and WTA (Hammack and Brown, 1974; Horowitz and McConnell, 2002) resulted in a number of experimental studies (see, e.g., Bateman et al, 1997; Kahneman et al., 1990) and theoretical model developments (see, e.g., Hanemann, 1991, 1999; Tversky and Kahneman, 1991). A second reason was the similarities of the experiments and valuation exercises, and the use of experimental methods to test the validity of hypothetical surveys (see, e.g., Carlsson and Martinsson, 2001; Cummings et al., 1995; Frykblom, 1997; Lusk and Schroeder, 2004; Neill et al., 1994). At the same time there are important philosophical differences between a standard behavioral economist and a standard stated preference economist. Exaggerating somewhat, we could say that the typical behavioral economist claims that preferences often are irrational, that they can be manipulated, and that it is not clear that preferences of the individual should be reflected in public policy. The typical stated preference economist takes preferences as given, even if they are irrational, and believes that they should not be manipulated and that preferences as expressed in the surveys are an important input for public policy. As we will discuss in this paper, there are a number of areas where stated preference can and should be developed in light of more recent findings in behavioral economics. The aim is therefore to discuss some areas within behavioral economics that are of interest for stated preferences in the sense that they can improve the reliability of our studies, and in particular we will discuss the role of stated preferences and the design of surveys. We will discuss four different areas: (i) incoherent preferences, (ii) learning and constructed preferences, (iii) context dependence, and (iv) hypothetical bias.

¹ Interestingly, a number of economists have made important contributions in both behavioral and environmental economics; four prominent examples are Glenn Harrison, Jack Knetsch, John List, and Jason Shogren. This is of course not to say that the work in behavioral economics has not faced any opposition in environmental economics.

2. Revealed and normative preferences

There is ample evidence in behavioral economics that people do not appear to do what is best for them. People smoke and drink too much, they do not study hard enough, they postpone writing the reviews until way past the deadline, and they stick with the default option even though it may not be the best option (see, e.g., Choi et al., 2003, 2004; Laibson, 1997). This means that it could be important to distinguish between revealed preferences and normative preferences (see, e.g., Beshears et al., 2008). Revealed preferences rationalize the individual's observed choices/decisions, while normative preferences represent the individual's actual interests. In many cases, revealed preferences should be interpreted as normative preferences, but due to for example decision-making errors, revealed preferences will not always represent normative preferences. Note that what we obtain in a stated preference (SP) survey are the revealed preferences. Beshears et al. (2008) discuss five factors that can create a wedge between revealed and normative preferences: (i) limited personal experience, (ii) complexity, (iii) passive choice/defaults, (iv) third-party marketing, and (iv) intertemporal choice. The first three are of obvious interest for SP surveys. There is evidence that people stick with default options even if they know it is not optimal, or that they stick with the default when it is set by others because they think the default was chosen for good reasons (see, e.g., Choi et al., 2004; Madrian and Shea, 2001). Complexity of the choice situation can have a number of effects on individuals, such as making them more likely to accept default options (O'Donoghue and Rabin, 1999), to make more errors (de Palma et al., 1994), or to adopt heuristic decision rules (Heiner, 1983). Finally, when it comes to limited personal experience, there is experimental evidence that experienced subjects suffer less from anomalies (List, 2003). Later on we will also discuss the literature on experienced versus decision utility (Kahneman et al., 1997).

Are there any implications of the above discussion for the design of SP surveys? One is to look at the difference between experienced and inexperienced respondents. Perhaps the revealed preferences in the survey situation of experienced respondents are less inconsistent with their normative preferences. Maybe experienced respondents make fewer errors when responding. However, there are large potential problems with endogeneity when comparing experience and inexperienced respondents. For example, a difference in willingness to pay for environmental conservation between respondents who have a lot of experience with say outdoor recreation and those who stay at home and read books is most likely not only due to differences in experience, but also differences in taste. One way of dealing with this problem

is to use exogenous events or differences between different samples. One such example is the study by Carlsson et al. (2009), who conducted a WTP study for avoiding power outages. In their study they had the possibility to conduct the SP study both before and after a large storm (although not with the same respondents). The storm caused power outages for around 20% of the Swedish households, and was covered in detail by the media. Thus, one might argue that the respondents who answered the survey after the storm had much better knowledge about power outages, and a sizeable proportion of the sample had direct experiences with a recent large outage. Interestingly, they found a lower WTP after the storm. In particular, there was a larger fraction of respondents with zero WTP. Consequently, in their particular case, experience resulted in a lower WTP.

The issue of task complexity has received considerable attention in the SP literature. Task complexity can potentially affect both the extent of inconsistent choices, the decision rules adopted by the respondents, and the welfare estimates (see, e.g., DeShazo and Fermo, 2002; Swait and Adamowicz, 2001). One interesting development is to try to reduce the task complexity, and one such example is virtual reality. Many SP surveys involve complex information, and it is difficult to experience the environment in the survey. Communicating public goods and risks with visual aids is of course nothing new (see, e.g., Carson et al, 2003; Corso et al., 2002), but virtual reality provides the respondent with much more freedom to explore different scenarios and to understand what would actually happen and how it would look like. There are two very recent papers that use virtual reality to communicate environmental changes: Fiore et al. (2009) and Bateman et al. (2009). For example, Bateman et al. use a virtual reality world in which respondents can fly around and explore the area. They import actual GIS data for an area, render the area as it looks today, and then simulate the change in the environment. One of the interesting findings is that the difference between willingness to pay for gains and willingness to accept the corresponding loss is smaller for the group of respondents facing the virtual reality treatment. Consequently, by reducing the task complexity with virtual reality, they manage to reduce the difference in willingness to pay and willingness to accept.

That people tend to stick with defaults is a similar type of behavior as when subjects stick with a status quo alternative. However, it is also a behavior that is similar to yea-saying, since one reason why people stick with defaults is that they believe that the default option has been designed by someone for a good reason. It is, however, unclear what the implications for the

design of SP surveys are. Whether or not to include an opt-out alternative mainly depends on the actual choice situation and what welfare measure the researchers want to measure.

Another way to think of incoherent preferences is to apply the concepts of decision and experienced utility (Kahneman et al., 1997; Kahneman and Sugden, 2005). Decision utility is what we study when we observe choices made by individuals. Experienced utility on the other hand is the utility that people experience for example at the time of consumption. Usually this also refers to a more hedonic measure of utility in terms of pain and pleasure. If people were rational, there would be no need to care about the concept of experience utility, or put differently the choices made based on decision utility would also maximize experience utility. However, there are several reasons for why decision and experience utility can deviate; see Kahneman and Sugden (2005) for a detailed review and discussion. The underlying reasons for the difference are of course similar to the ones explaining a difference between revealed and normative preferences. One important reason for a deviation is referred to as adaptation of the hedonic treadmill, which means that humans adapt quickly to changes. This means that a positive experience or happiness becomes less intense over time. There are several reasons for adaptation, for example changing standards of evaluation and redeployment of attention.

If we care about the utility of the actual experience or the normative preference, then it is problematic to use decision utility as an index of welfare. This has led to a growing literature on paternalism within behavioral economics that argues that if preferences are incoherent/irrational, then there is room for policy makers to use their own judgment about what is best for an individual. At the same time, the government should not unnecessarily interfere with the lives of individuals, something that has led to terms such as libertarian paternalism (Sunstein and Thaler, 2003a, 2003b) and regulation for conservatives (Camerer et al., 2003). For example, if people stick with the default option despite that option being the worst, there is room for a policy maker to affect the design of the default option. The literature on paternalism has faced a lot of opposition; see, e.g., Sugden (2007, 2008). One question is how to deal with the fact that people adapt quickly to negative and positive outcomes (Loewenstein and Ubel, 2008).

Using stated happiness or well-being questions is an alternative to SP surveys that would directly measure experience utility (see, e.g., Luechinger, 2009; van Praag and Baarsma,

2005; Welsch, 2009).² The idea is simple: the level of the public good/externality is correlated with individuals' reported subjective well-being. This way, a measure of the value of the public good is obtained in terms of life satisfaction/happiness, but it is also possible to express the value relative to the effect of income on happiness. There are a number of advantages with this approach compared to an SP survey. There are no direct incentives for respondents to overstate or understate their well-being, or at least no incentives that are related to the level of the public good. Happiness, or well-being, is one measure of experience utility. Hence, it might seem compelling to argue for an increased use of well-being questions in environmental economics, and indeed a number of authors do just that (see, e.g., Frey and Stutzer, 2002; Frey et al., 2004). However, there are also a number of serious disadvantages with the method that we need to be aware of, e.g., , that we do not know what information subjects have about the good, that the method is actually rather data intensive, and that for many environmental problems it is by definition difficult to ex ante obtain information about experience utility. Even if this method has a number of disadvantages, it would be interesting to conduct SP studies and compare them with well-being studies. In particular, it would be interesting to investigate whether they really give significantly different results, and if so under what circumstances.

3. Learning and constructed preferences

It is important to distinguish between incoherent preferences and learning in for example an SP survey. If a respondent does not have stable preferences throughout a choice experiment, it does not have to imply that he or she is making inconsistent choices due to for example decision errors. Instead, the respondent could be learning his or her preferences. We know that participating in an SP survey is not easy. Subjects receive a lot of information, many times on things that they are unfamiliar with. Then we create a "market," and ask them to make choices in this market. It is thus rather likely that some, or many, of the respondents do not have a clear picture of what their preferences are. This means that respondents could be forming and even changing their preferences while answering the survey. As argued by Plott (1996), stable and theoretically consistent preferences are the product of experience gained through practice and repetition. Practice and repetition could take place in the marketplace, but also in the survey situation; as shown by Bateman et al. (2008), respondents might learn the institutional design by responding to several double-bounded CVM (Contingent Valuation

² An example of a question is "On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?"

Method) questions. There is also evidence that repeated behavior reduces anomalies and in particular that more experienced traders are less inconsistent (List, 2003). These findings have two implications when looking at responses in an SP survey: (i) preferences might seem incoherent, but they are not, and (ii) preferences elicited at a later stage in the survey instrument are less noisy and better reflect the respondent's normative preferences. In survey formats with repeated questions there are thus reasons to include warm-up questions or simply ignore the responses to the first set of questions.

If learning and construction of preferences is common in SP surveys, it could have implications for the choice of question format. For example, with CVM single-bounded questions, respondents only get one shot at expressing their preferences, while with other formats such as bidding games and choice experiments, respondents make repeated choices. If we conduct an SP survey on a good involving attributes that are not very familiar to the respondent, there is a risk that his/her preferences are not really formed before the survey situation. If we then conduct a test of the stability of the preferences, we might find that the preferences are not stable.³ It is then not obvious that this is related to problems with the method itself; instead it could be that preferences are constructed as the respondent goes through the survey. Note the reverse argument here: if we find that preferences are not stable across the choice sets in a choice experiment, it could be an argument for using a choice experiment, but focusing on the responses to the later questions. However, there is a potential counter effect to learning referred to as coherent arbitrariness (Ariely et al., 2008). This means that individuals' choices are often internally coherent, but at the same time they can be strongly anchored to some initial starting point. The manifestation of this behavior in SP surveys is starting-point-bias (Herriges and Shogren, 1996; Ladenburg and Olsen, 2008). Consequently, if we wish to investigate learning in SP surveys we need to be able to control for coherent arbitrariness. A recent and excellent example of such a study is Bateman et al. (2008). They devise experiments in order to test three conceptions of individuals' preferences as (i) a-priori-well-formed, (ii) discovered or learned through repetition, or (iii) internally coherent but arbitrary. They conduct a sequence of double-bounded CVM questions (on a set of animal welfare-related goods). The initial bid presented in the single-bounded question for a given good is then an arbitrary stimulus. If there is coherent arbitrariness, then the initial bid should influence behavior in the subsequent valuation tasks. Their main findings are that (i)

³ The evidence on stability of preferences is mixed; see, e.g., Johnson et al. (2000), Carlsson and Martinsson (2001), and Layton and Brown (2000).

there is institutional learning, i.e., increasing familiarity with the contingent market results in greater consistency of valuation responses between the single-bounded and double-bounded formats, and (ii) there is anchoring on initial bids, but it decreases across valuation tasks. Consequently, inconsistencies and anomalies are reduced as repeated valuation tasks are performed.

4. The importance of context dependence

One important message to take home from not only experimental economics but also valuation studies is that the context matters. This is of course nothing new for most valuation practitioners. What we have learned from behavioral economics however is that context dependence is not unique to SP methods or to environmental goods. We know that things like whether the behavior is observed by others or not, and how the problem is framed will affect respondent behavior both in surveys, in experiments, and in field settings. Context is of course a broad concept, but includes factors such as whether the action is observed by others (e.g., Alpizar et al. 2008a, Andreoni and Petri, 2004; Legget et al., 2003; List et al., 2004), information about what others do (Bardsley and Sausgruber, 2005; Carlsson et al., 2008; Frey and Meier, 2004; Shang and Croson, 2006), and the characteristics of the interviewer/solicitor (Bateman and Mawby, 2004; Landry et al., 2006).

One interesting question is whether respondent behavior is more sensitive to context (for example in terms of perception of the behaviors of others) when making a hypothetical (but realistic) choice than when making a choice that involves an actual payment. Some suggest that the difference may be large (Bertrand and Mullainathan, 2001) while others say it is small (Hanemann, 1994). Factors such as unfamiliarity, complex information, and public good character might make environmental valuation more prone to anomalies, but we know that they occur in other settings as well. Using a natural field experiment, Alpizar et al. (2008b) investigated the importance of the social context for people's voluntary contributions to a national park in Costa Rica. Some subjects made actual contributions while others stated their hypothetical contribution. Both the degree of anonymity and information provided about the contributions of others influenced subject contributions in the hypothesized direction. They found a substantial hypothetical bias with regard to the amount contributed. However, the influence of the social contexts is about the same when the subjects made actual monetary contributions as when they stated their hypothetical contributions, suggesting that SP methods

can be used to correctly infer about contextual effects. Clearly, more studies comparing the effects of context in hypothetical and real settings are needed in order to confirm this finding.

Actually, the tools we have are highly suitable for analyzing context dependence. The attribute-based random utility model can readily incorporate contextual elements. Our survey instruments can easily be adapted to different contexts, and it is rather easy to design different treatments with different contexts (Swait et al., 2002). Two examples of such studies are Hu et al. (2006), who study the effect of labels and reference points on food demand, and List et al. (2004), who study the effect of the extent of social isolation on hypothetical charitable contributions.

The importance of context dependence also have important implications for validity testing of SP methods: a comparison between hypothetical and actual behavior should be done for a given social context. This is something we will come back to in the next section, when we discuss hypothetical bias. There are also implications for pooling of data from different sources. For example, combining results from SP and travel cost studies could be problematic due to large contextual differences.

The main goal of valuation studies is to provide information for some kind of welfare analysis of a non-market good. The context dependence of valuation studies means that we should be careful when using the results in a welfare analysis. At the same time, experimental evidence suggests that actual behavior for example in the market is also vulnerable to the context. So again, context dependence is not a problem of SP only, but rather a potential problem for welfare evaluations in general; see, e.g., Tversky and Simonson (1993) and Shogren and Taylor (2008). Whether it poses a greater problem for SP than for revealed preferences still remains an open question, but some experimental evidence suggests that it might not.

5. Hypothetical bias

There has been an extensive discussion about the possibility of eliciting preferences, both for private and public goods, with SP methods and the extent of hypothetical bias; see, e.g., List and Gallet (2001) and Murphy et al. (2005). With hypothetical bias we mean the bias introduced by asking a hypothetical question and not confronting the respondent with a real situation.

There are a number of cases where laboratory and natural field experiments can be used as an alternative to SP surveys. The main application is perhaps within food economics, where both SP surveys and experiments have been used to elicit preferences for food attributes and the effects of information on consumer choices (see, e.g., Noussair et al., 2004; Lusk et al., 2006). A number of studies compare results from lab experiments with SP surveys on the same topic. Most of these are done with the purpose of testing for hypothetical bias. The results are clearly mixed. Carlsson and Martinsson (2001) failed to reject a hypothesis of equal marginal WTP in a real and a hypothetical setting (both conducted in a lab), while Johansson-Stenman and Svedsäter (2008) did reject the equality of marginal WTPs and Lusk and Schroeder (2004) found that hypothetical choices overestimate total WTP, although they could not reject the equality of marginal WTPs for changes in individual attributes. Then there are studies that compare lab experiments, SP surveys, and behaviors outside the lab. Shogren et al. (1999) conducted a hypothetical mail survey and a lab experiment concerning irradiated food, and compared the results with actual store purchases. They found that both the survey and the lab experiment resulted in a larger predicted market share of irradiated chicken than the actual market share in the grocery store. Chang et al. (2009) found that both an SP survey and an actual lab experiment predict actual retail sales fairly well, although the non-hypothetical experiment performs better than the hypothetical choice experiment. The most interesting finding is perhaps the one by Lusk et al. (2006), who compared a framed field experiment with actual retail sales. They found that the results of the framed field experiment predicted consumer behavior in the store, although there was some evidence of more pro-social behavior in the framed field experiment.

There are two important lessons from the above discussion. First, we should be careful when comparing the results obtained from either a laboratory experiment or an SP survey to actual behavior. Clearly, differences cannot only be explained by hypothetical bias in the strictest sense, since both actual lab experiments and SP survey results have some problems predicting actual retail behavior. Second, although the results above are mixed, actual lab experiments seem to more often perform better than SP studies.

As discussed by Levitt and List (2007), a number of factors can explain the behavioral differences between the laboratory and the real world: scrutiny, context, stakes, selection of subjects, and restrictions on time horizons and choice sets. Some of these can of course also explain the differences between actual behavior and surveys and lab experiments. The first is

the degree of scrutiny. In both the lab and the survey situation, subjects/respondents know that they are taking part in a study where someone is interested in their behavior. The second is the choice-set restriction. In the lab and the survey, the choice sets are clearly defined and restricted, while in a store the choice-sets are larger and perhaps less clear. The third is the context in which the choices are made, i.e., the store versus the lab. The vast literature comparing behavior in the lab and the field shows that these could be important factors explaining the difference. For example, the degree of anonymity in experiments is one potential measure of scrutiny, and a number of experiments show that the extent of pro-social behavior increases the less anonymous the decisions (see, e.g., List et al., 2004; Rege and Telle, 2004; Soetevent, 2005). The effects of choice-set restrictions can also be important. For example, Bardsely (2008) and List (2007) found that subject behavior in a traditional dictator game changes when there is a possibility of taking money from the recipient's endowment.

Clearly, it would be beneficial to carefully consider differences between, say, behavior in a survey and behavior involving actual money in the same setting as in Levitt and List (2007). Let us discuss a few examples. One important difference between a survey situation and an actual decision is the focus on the issue at hand in the survey situation. This is nothing new for survey practitioners, and for example the suggestion to remind subjects about the budget constraint is one way to try to widen the focus of the respondents. Another interesting way of addressing this, which has great potential, is the time-to-think protocol (Whittington et al., 1992; Cook et al., 2007). The idea is very simple, and it can be done in several different ways. The design is similar to a standard mail survey. For example, an enumerator can come to a respondent's home and introduce and explain the survey, but the respondent is not asked to answer the survey immediately. Instead the enumerator leaves, and the respondent gets time to think about the question and can answer for example the next day. Surveys can then be collected, or sent back by mail. In the study by Cook et al. (2007), respondents made fewer inconsistent responses in a treatment where they were given time to think compared to a standard treatment, and the WTP was also significantly lower. There are a number of features of this design that can reduce hypothetical bias, and several of them have to do with contextual differences between the survey situation and the corresponding actual situation. First of all, an interviewer is not present when the decisions are made. Second, by giving time to think, it is likely that the choice set restriction of the survey is less restricting, and the focus on the scenario is reduced. Third, respondents are allowed to talk with other household members, friends, etc. about the issue. This is likely to better resemble an actual situation. At

the same time, there are a number of potential problems with this protocol. For example, it is not at all clear *whose* preferences are elicited. Since we cannot control who the respondent talks to, another household member or a friend could strongly influence the responses. However, this does not always have to be something undesirable. First of all, we might be interested in the response of the household and not an individual member. Second, this could actually better reflect how decisions are made.

One important message from the above discussion is that it is very difficult to test for hypothetical bias. Since it is likely that the context affects behavior, most comparisons between a real situation and a survey situation would imply a number of contextual differences. The simple fact that most SP studies involve an interviewer while most real situations do not should be enough to make this point. So, one alternative then is to reduce the contextual differences to a minimum. This means for example comparing lab experiments with SP surveys, but as discussed above, there is evidence that also lab experiments have problems with predicting behavior outside the lab.

There is one aspect of respondent behavior and hypothetical bias that has hardly received any attention as far as we know, and that is aversion to lying. In an experimental setting it has been shown that *some* people have an aversion to lying (see, e.g., Gneezy, 2005). Gneezy (2005) used a sender-receiver game designed such that a person who lies benefits from it while the other subjects' payoff decreases. Two important results arise from this study. The first is that not everybody lies. The second is that the probability of lying increases in the gain, and decreases in the loss of the other subject, i.e. the more other are hurt, the less likely I am to lie. What this suggests is that respondents, even if they understand that it is in their interest to act strategically, do not do so because they do not like to lie. One option is to try to identify subjects, or characteristics of subjects, with a strong lying aversion. Another is to increase the likelihood of lying aversion among the subjects. Right now, we do not know whether either of these options is possible to do. One hypothesis is that the degree of scrutiny and the sense of accountability increases lying aversion. We could tell subjects that we will ask them a number of questions about their answers, ask them to justify them, and ask them to explain to us how they would accommodate their responses within their budget.

6. Conclusions

There is a lot to learn from behavioral economics when it comes to designing and conducting stated preference surveys. We have in this paper discussed a few areas where the design of stated preferences could be improved. At the same time we argue that stated preference methods are suitable for analyzing many of the issues that interest behavioral economists, such as context dependence and learning.

References

- Alpizar, F., F. Carlsson and O. Johansson-Stenman (2008a) Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics* 92, 1047-1060.
- Alpizar, F., F. Carlsson and O. Johansson-Stenman (2008b) Does context matter more for hypothetical than for actual contributions? Evidence from a natural field experiment. *Experimental Economics* 11, 299-314.
- Andreoni, J., Petrie, R., (2004) Public good experiments without confidentiality: a glimpse into fund-raising, *Journal of Public Economics* 88, 1605–1623.
- Ariely, D., G. Loewenstein and D. Prelec (2008) Coherent arbitrariness: Stable demand curves without stable preferences, Working Paper Massachusetts Institute of Technology.
- Bardsley, N. (2008) Dictator game giving: Altruism or artifact? *Experimental Economics*, 11, 122-133.
- Bardsley, N. and R. Sausgruber (2005) Conformity and reciprocity in public good provision. *Journal of Economic Psychology* 26, 664–681.
- Bateman, I. D. Burgess, G. Hutchinson, D. Matthews (2008) Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. *Journal of Environmental Economics and Management* 55, 127-141.
- Bateman, I., Day, B., Jones, A., Jude, S. (2009) Reducing gain–loss asymmetry: A virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management* 58, 106-118.
- Bateman, I. and J. Mawby (2004) First impressions count: interviewer appearance and information effects in stated preference studies. *Ecological Economics* 49, 47-55.
- Bateman, I. A. Munro, B. Rhodes, C. Starmer and R. Sugden (1997) A Test of the Theory of Reference-Dependent Preferences. *Quarterly Journal of Economics* 112, 479-505.
- Beshears, J., J. Choi, D. Laibson and B. Madrian (2008) How are preferences revealed? Working paper.
- Bertrand, M. and S. Mullainathan (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review, Papers and Proceedings*, 91, 67–72.
- Camerer, C., Issacharoff, S., Loewenstein, G., O’Donoghue, T., and Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for ‘asymmetric paternalism’. *University of Pennsylvania Law Review*, 151, 1211–1254.

- Carlsson, F., J. Garcia and Å. Löfgren (2008) Conformity and the demand of environmental goods, Working Papers in Economics No. 286, Department of Economics, University of Gothenburg.
- Carlsson, F. and P. Martinsson (2001) Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments? *Journal of Environmental Economics and Management* 41: 179-92.
- Carlsson, F., P. Martinsson and A. Akay (2009) The effect of power outages and cheap talk on willingness to pay to reduce outages, IZA Discussion paper series No. 4307.
- Carson, R., R. Mitchell, M. Hanemann, R. Kopp, S. Presser, P.A. Ruud (2003) Contingent valuation and lost passive use: damages from the Exxon Valdez oil spill. *Environmental and Resource Economics* 25, 257–286.
- Chang, J.B., J. Lusk and F.B. Norwood (2009) How closely do hypothetical surveys and laboratory experiments predict field behavior? *American Journal of Agricultural Economics* 91, 518-534.
- Choi, J., D. Laibson, B. Madrian and A. Metrick (2003) Optimal defaults. *American Economic Review* 93, Papers and Proceedings 180-185.
- Choi, J., D. Laibson, and B.C. Madrian (2004) Plan Design and 401(k) savings outcomes. *National Tax Journal* 57, 275–298.
- Cook, J., D. Whittington, D. Canh, F.R. Johnson and A. Nyamete (2007) Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue Vietnam. *Economic Inquiry* 45, 100-114.
- Corso, P. J. Hammitt and J. Graham (2002) Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation. *Journal of Risk and Uncertainty* 23, 165-184.
- Cummings, R., Harrison, G. and E. Rutström (1995) Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive- Compatible? *The American Economic Review* 85, 260-266.
- De Shazo, J.R. and G. Fermo (2002) Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency, *Journal of Environmental Economics and Management* 44, 123–43.
- de Palma, A., G.M. Myers, and Y.Y. Papageorgiou (1995) Rational choice under an imperfect ability to choose. *American Economic Review* 84, 419-440.
- Fiore, S., G. Harrison, C. Hughes, E. Rutström (2009) Virtual experiments and environmental policy. *Journal of Environmental Economics and Management* 57, 65-86.

- Frey, B., S. Luechinger and A. Stutzer (2004) Valuing Public Goods: The Life Satisfaction Approach. CESifo Working Paper Series No. 1158.
- Frey, B. and S. Meier (2004) Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *American Economic Review* 94, 1717–1722.
- Frey, B. and A. Stutzer (2002) What Can Economists Learn from Happiness Research? *Journal of Economic Literature* 40, 402-435.
- Frykblom, P. (1997) Hypothetical question modes and real willingness to pay. *Journal of Environmental Economics and Management* 34, 275-287.
- Gneezy, U. (2005) Deception: the role of consequences. *American Economic Review* 95, 384-394.
- Hammack, J. and G. Brown (1974) *Waterfowl and wetlands: Toward bio-economic analysis*, Baltimore: John-Hopkins Press.
- Hanemann, M. (1991) Willingness to pay and willingness to accept: How much can they differ? *American Economic Review* 81, 635-647.
- Hanemann, M. (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives*, 8, 19–43.
- Hanemann, M. (1999) The economic theory of WTP and WTA, in “Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries” Bateman and K. G. Willis, Eds., Oxford Univ. Press, London.
- Heiner, R.A. (1983) The origin of predictable behavior. *American Economic Review* 73, 560-595.
- Herriges, J. and J. Shogren (1996) Starting point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management* 30, 112–131.
- Horowitz, J. and K. McConnell (2002) A review of WTA/WTP studies, *Journal of Environmental Economics and Management* 44, 426-447.
- Hu, W., W. Adamowicz and M. Veeman (2006) Labeling context and reference point effects in models of food attribute demand. *American Journal of Agricultural Economics* 88, 1034-1049.
- Johansson-Stenman O. and H. Svedsäter (2008) Measuring hypothetical bias in choice experiments: The importance of cognitive consistency. *B-E Journal of Economic Analysis & Policy* 8, Article 41

- Johnson R., W. Matthews and M. Bingham (2000) Evaluating welfare-theoretic consistency in multiple response, stated-preference survey. TER Working Paper T-0003.
- Kahneman, D., J. Knetsch and R. Thaler (1990) Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy* 98, 1325-1348.
- Kahneman, D. and R. Sugden (2005) Experienced utility as a standard of policy evaluation, *Environmental and Resource Economics* 32, 161-181.
- Kahneman, D., P. Wakker and R. Sarin (1997) Back to Bentham: explorations of experienced utility. *Quarterly Journal of Economics* 112, 375–405.
- Ladenburg, J. and S. Olsen (2008) Gender specific starting point bias in choice experiments: Evidence from an empirical study. *Journal of Environmental Economics and Management*, 56, 275-285.
- Laibson, D. (1997) Golden eggs and hyperbolic discounting, *Quarterly Journal of Economics* 112, 443-477.
- Landry, C., A. Lange, J. List, M. Price and N. Rupp (2006) Toward an understanding of the economics of charity: evidence from a field experiment. *Quarterly Journal of Economics* 121, 747–782.
- Layton, D. and G. Brown (2000) Heterogenous preferences regarding global climate change. *Review of Economics and Statistics* 82, 616-624.
- Legget, C. N., Kleckner, K. Boyle, J. Duffield and R. Mitchell (2003). Social desirability bias in contingent valuation surveys administered through in-person interviews. *Land Economics*, 79, 561–575.
- Levitt, S. D., and J.A. List, (2007) What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21, 153–74.
- List, J. (2003) Does market experience eliminate market anomalies? *Quarterly Journal of Economics* 118, 41–72.
- List, J.A., (2007) On the interpretation of giving in dictator games. *Journal of Political Economy* 115, 482-493.
- List, J., R. Berrens, A. Bohara and J. Kerkvilet (2004) Examining the role of social isolation on stated preferences. *American Economic Review* 94, 741–752.
- List, J. and C. Gallet (2001) What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values? *Environmental and Resource Economics* 20, 241-254.
- Loewenstein, G. and P. Ubel (2008) Hedonic adaptation and the role of decision experience utility in public policy. *Journal of Public Economics* 92, 1795-1810.

- Luechinger, S. (2009) Valuing air quality using the life satisfaction approach. *Economic Journal* 119, 482-515.
- Lusk, J., J. Pruitt and B. Norwood (2006) External validity of a framed field experiment. *Economics Letters* 93, 285-290.
- Lusk, J. and T.C. Schroeder. (2004) Are choice experiments incentive compatible? A test with quality differentiated beefsteaks. *American Journal of Agricultural Economics* 85, 840-856.
- Madrian, B.C. and D.F. Shea (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics* 116, 1149-1187.
- Murphy, J. G. Allen, T. Stevens and D. Weatherhead (2005) A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics* 30, 313-325.
- Neill, H. R. Cummings, P. Ganderton, G. Harrison, and T. McGuckin (1994) Hypothetical surveys and real economic commitments. *Land Economics* 70, 145-154.
- Noussair, C., S. Robin and B. Ruffieux (2004) Do consumers really refuse to buy genetically modified food. *Economic Journal* 114, 102-120.
- O'Donoghue, T. and M. Rabin (1999) Doing it now or later, *American Economic Review* 89, 103-124.
- Plott, C. (1996) Rational individual behavior in markets and social choice processes: the discovered preference hypothesis, in: K. Arrow, E. Colombatto, M. Perleman, C. Schmidt (Eds.), *Rational Foundations of Economic Behavior*, Macmillan, London, St. Martin's, New York, pp. 225–250.
- van Praag, B.M.S. and B. Baarsma (2005) Using happiness surveys to value intangibles: the case of airport noise. *Economic Journal* 115, 224–246.
- Rege, M. and K. Telle (2004) The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics* 88, 1625–1644.
- Shang, J. and R. Croson (2006) Field experiments in charitable contribution: The impact of social influence on the voluntary provision of public goods. Working Paper.
- Shogren, J.F., J.A. Fox, D.J. Hayes, and J. Roosen (1999) Observed for Food Safety in Retail, Survey, and Auction Markets. *American Journal of Agricultural Economics* 81, 1192–99.
- Shogren, J. and L. Taylor (2008) On behavioral-environmental economics. *Review of Environmental Economics and Policy* 2, 26-44.

- Soetevent, A.R. (2005) Anonymity in giving in a natural context: An economic field experiment in thirty churches. *Journal of Public Economics* 89, 2301–2323.
- Sunstein, C., and R. Thaler (2003a) Libertarian paternalism. *American Economic Review, Papers and Proceedings*, 93(2), 175–179.
- Sunstein, C., and R. Thaler (2003b) Libertarian paternalism is not an oxymoron. *University of Chicago Law Review*, 70, 1159–1202.
- Sugden, R. (2007) Cost-Benefit Analysis as Market Simulation. A new approach to the problem of anomalies in environmental valuation. RFF Discussion Paper 07-28, Resources for the Future, Washington D.C.
- Sugden, R. (2008) Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19, 226-248.
- Swait, J. and W. Adamowicz (2001) The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. *Journal of Consumer Research* 28, 135-148.
- Swait, J., W. Adamowicz, M. Hanemann, A. Diederich, J. Krosnick, D. Layton, W. Provencher, D. Schkade, and R. Tourangeau (2002) Context Dependence and Aggregation in Disaggregate Choice Analysis. *Marketing Letters* 13, 195-205.
- Tversky, A. and D. Kahneman (1991) Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics* 106, 1039-1061
- Tversky, A. and I. Simonson (1993) Context-dependent preferences. *Management Science* 39, 1179-1189.
- Welsch, H. (2009) Implications of happiness research for environmental economics, *Ecological Economics* 68, 2735-2742.
- Whittington, D., V. K. Smith, A. Okorafor, A. Okore, J. L. Liu, and A. McPhail (1992) Giving Respondents Time to Think in Contingent Valuation Studies: A Developing Country Application. *Journal of Environmental Economics and Management* 22, 1992, 205–25.