When Are Preferences Consistent? The Effects of Task Familiarity and Contextual Cues on Revealed and Stated Preferences

Felix Schläpfer, Baruch Fischhoff

August 2010
When Are Preferences Consistent? The Effects of Task Familiarity and Contextual Cues on Revealed and Stated Preferences

August 2010

Author’s address: Felix Schläpfer
E-mail: felix.schlaepfer@soi.uzh.ch

Baruch Fischhoff
E.mail: baruch@cmu.edu
When Are Preferences Consistent?
The Effects of Task Familiarity and Contextual Cues
on Revealed and Stated Preferences

Felix Schläpfer1* and Baruch Fischhoff2

1Department of Economics, University of Zurich and
Institute for Environmental Decisions, Federal Institute of Technology ETH
Zurich; Switzerland

2Department of Social and Decision Sciences
Department of Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213-3890; USA
Tel: +1-412-268-3246; email: baruch@cmu.edu

*Address for correspondence: Department of Economics, University of Zurich,
Hottingerstr. 10, 8092 Zurich; Switzerland, E-mail: felix.schlaepfer@soi.uzh.ch.

June 2010

Abstract

Traditionally, economists make a sharp distinction between stated and revealed preferences, viewing the latter as more fully meeting the assumptions of economic analysis. Here, we consider one form of empirical evidence regarding this belief: the consistency of choices in stated and revealed preference tasks. We show that both kinds of task can produce consistent choices, suggesting that both can measure underlying preferences, if necessary conditions are met. We propose that a necessary condition is that task be either familiar to those facing it or offer contextual cues that substitute for familiarity, such as prices in competitive markets or recommendations from trusted, knowledgeable sources. We show that how well decision makers achieve such understanding is often confounded with the method that researchers use. Considering task familiarity not only clarifies some of the conflicting evidence regarding revealed and stated preference methods, but raises potentially productive questions regarding the roles of social institutions in shaping preferences.

Keywords: Consistency, contingent valuation, framing, public goods, revealed preferences, stated preferences, validity.

JEL codes: D01, D03, Q51

We are grateful to James Murphy and his co-authors for making the dataset of their meta-analysis available.
1. Introduction

Economic analysis traditionally relies on revealed choice, arguing that, for practical purposes, individuals’ preferences meet standard economic assumptions (e.g., Friedman 1953). Economists have maintained this position despite growing challenges (e.g., Simon 1956; Conlisk 1996; McFadden 2001; Ariely et al., 2003; Kahneman, 2003; Camerer and Fehr, 2006), partly because they lack an account of when these violations of its assumptions matter and partly because they distrust the alternative, stated preference methods. On theoretical grounds, many economists are wary of stated preferences because they know how, in theory, individuals could respond strategically to questions about alternative levels of public good provision (Gibbard, 1973; Satterthwaite, 1975). On empirical grounds, many economists note the non-rational behavior first identified in stated preference studies (e.g., Kahneman and Tversky, 1979).

Despite the strength of this opposition, the evidence against stated preference methods is fairly limited (Dominitz and Manski 1997). Many anomalies have been replicated in tasks with monetary stakes, indicating analogous problems with at least some revealed preferences (Camerer and Hogarth, 1999). Moreover, the largest body of evidence suggesting problems with stated preferences comes from the somewhat anomalous context of contingent valuation surveys (Arrow et al. 1993; Mitchell and Carson, 1989), in which respondents state their willingness to pay for public goods that are far from their personal experience and have not been publicly discussed. When those tasks are described in simple terms, they force respondents to figure out what exactly they are evaluating (Fischhoff and Furby, 1988). When the tasks are described in specific terms, they pose cognitive demands on respondents. It is hard for individuals to respond consistently when they do not know what question they are answering or cannot keep all its details in mind.

A continuing debate among dedicated scientists suggests that there is no single, general answer to the question of the relative validity of revealed and stated preference methods. However, there may be answers to where or when each is more valid (Fischhoff 1991, p. 846; Slembeck and Tyran 2004, p. 683; Druckman 2004, p. 683; Shogren 2006, p. 169). Here, we examine the cognitive tractability of revealed and stated preference tasks selected to vary on three features of the choice task: (a) whether the good is familiar, (b) whether informative contextual cues accompany the task, and (c) whether the good is private or public.

Section 2 defines our criterion for consistent preferences. Section 3 presents the three features of the choice task. Section 4 examines selected studies representing the different possible combinations of these features, in terms of the consistency of the stated and revealed preferences that they produce. Section 5 reports statistical analyses of the roles of these features in predicting one form of consistency in a large set of studies. Section 6 concludes with a proposal for viewing revealed and stated preferences as complementary, rather than competing approaches, illuminating different aspects of preference-forming processes.
2. Empirical consistency criteria

Economic theory and analysis are based on the axiomatic assumptions that preferences are complete and transitive. Preferences that satisfied these assumptions should demonstrate the kind of consistency that psychologists call *construct validity*. They should be (1) insensitive to irrelevant information, and (2) appropriately sensitive to relevant information.

A well-known violation of the first condition arises with anchoring effects, in which responses are affected by arbitrary numbers (e.g., think of the last two digits of your Social Security number, before answering). An example is the anchoring effect in responses to hypothetical referendum questions in contingent valuation surveys with “alternative bid designs”: The binary (yes/no) responses express lower values when the question is, “Would you pay [$1/$2/$3, etc]?” than when it is, “Would you pay [$10/$20/$30, etc.]?” (e.g. McFadden 2001). An anchor can also convey information, such as what researchers believe the value to be. If so, then responses should be sensitive to it.

It is harder to establish whether sensitivity to relevant information is appropriate. *Scope* tests ask whether preferences increase with increases in the quantity of a good. However, these are vague tests, without specifying how great the sensitivity should be. Moreover, subadditivity, as in the finding that 57 wilderness areas were valued only about fifty percent more than one wilderness area (McFadden, 1994), has been explained in terms of “substitution effects and diminishing marginal rates of substitution” (Hahneman’s (1994, p. 35). Hahneman’s argument implies that respondents believe that the 57 wilderness areas are almost indistinguishable (hence are substitutable) and that there is little extra value in having an additional 56 (diminishing marginal utility).

Thus, construct validity entails a theoretical argument, regarding which features matter to people (hence should evoke sensitivity) and how methodologically sound studies are. Below, we outline construct validity arguments for the eight classes of study mentioned above (i.e., all combinations of familiar and unfamiliar, public and private goods, with and without informative contextual cues).

We restrict the discussion to studies with *between-subject* designs, where different respondents evaluate, say, different amounts of a good. *Within-subject* designs, where each respondent evaluates multiple versions of a good, inevitably provide hints as to what kinds of consistency researchers expect (Poulton, 1968, 1994; Frederick and Fischhoff 1998, Amir and Levav 2008).

3. Classification of choice domains

In the analysis below, we use the following definitions:

(1) *familiarity*: extent of experience with the good (e.g., the proposed change of the specific wilderness areas).

(2) *informativeness of contextual cues*: how well they convey the issues relevant to the choice, either directly or through surrogates, such as market prices, medical advice, or public debates.
(3) private vs. public goods: “private goods” are for personal consumption; “public goods” are public services or amenities, as might be evaluated with votes in hypothetical referenda or on rating scales.

“Revealed preferences” (RP) refers to consequential choices. With private goods, they involve actual transactions. With public goods, they require implementation rules, such as a simple majority voting rule with a coercive payment mechanism.

“Stated preferences” (SP) refers to hypothetical choices without a clear, coercive implementation rule.

Table 1 summarizes the 2x2x2x2 design of the analyses that follow.

[Table 1 about here]

4. Exploratory analysis of consistency across choice domains

4.1. Familiar private goods with informative contextual cues

Many marketing studies and observational studies of consumer behavior have these properties. Taken together, they provide overwhelming evidence of price sensitivity, an essential form of choice consistency, in the preferences revealed in actual markets for gasoline, computers, air trips and many other goods. Experimental studies of choice consistency, in such domains are inherently rare because natural choice sets and cues (e.g., equilibrium prices, advertisements, consumer advice) cannot be manipulated without creating unnatural cues that might elicit unrepresentative behavior. For example, manipulating actual prices can send misleading (or just confusing) signals about quality, producing inconsistent choices. Conversely, consistent choices can be overdetermined, potentially reflecting habitual responses, contextual cues, or social pressure, as well as tapping robust preferences (List 2004).

Revealed preferences. In one well-known experimental test, with a familiar, private choice, and clear (contextual) price information, Dickie et al. (1987) offered boxes of strawberries door-to-door at different prices. The resulting demand curve suggests sensible sensitivity to price information, perhaps reflecting personal experience with the good, perhaps reflecting contextual cues from prices in ordinary markets.

Stated preferences. Dickie et al. (1987) also asked other people the hypothetical question of how many boxes they would buy at the prices offered in the revealed preference portion of their study. The hypothetical choices produced a demand curve similar to that for the actual ones.

Indirect evidence of choice consistency with familiar private goods and informative contextual cues is seen in the weaker endowment effect found with experienced traders in market settings, compared to that found with inexperienced traders, unfamiliar with the good and market context (List 2004).

4.2. Private goods, directly experienced but without informative contextual cues

Strong experimental tests of construct validity, for such choices, are difficult because it is hard to remove natural context cues (e.g., competitive market prices) without creating an unnatural choice, potentially inducing unrepresentative behavior.
Revealed preferences. Removing familiar contexts is, however, a staple of marketers, hoping to entice consumers into making choices inconsistent with their normal preferences. To this end, casinos use chips (rather than money), have no clocks, and ply gamblers with drinks. Gamblers’ remorse is a special, sometimes tragic case of inconsistent choices in revealed preferences.

Stated preferences. Experiments involving familiar private goods, but without their normal informational context are also a staple of psychological and behavioural economics researchers, seeking inconsistent choices that will demonstrate inappropriate sensitivity to contextual cues. Loomis et al. (1997), however, showed cases in which choices regarding a fairly familiar good (a wildlife art print) offered in unusual contexts were sensitive to experimental variations in price. In a demonstration of the power of the cues provided by within-subject designs, Frederick and Fischhoff (1998) found that stated willingness to pay for toilet paper was 2.5 times more sensitive to quantity when it was manipulated within rather than between subjects, indicating either that the former choices were too sensitive or that the latter were insufficiently sensitive (Frederick and Fischhoff 1998). Levin and Gaeth (1988) found a case in which great familiarity with a product (ground beef) overcame an experimental attempt to manipulate contextual cues – seen in failure to find a familiar framing effect.

4.3. Private goods, without direct experience, but with informative contextual cues

Here, although individuals are unfamiliar with the goods themselves, they have access to useful cues, for instance from market context or from a trustworthy source with enough information to make consistent choices possible. That source might be a knowledgeable relative, consumer organization, or government regulator. The source must know enough about the decision makers to identify the facts (and uncertainties) most relevant to their choices. The relevant knowledge might include predicting how those individuals will experience novel goods. Taking advantage of such sources requires being able to assess their knowledge and honesty. Many investors feel that they have misplaced their trust in financial advisors, who have proven to lack one property or the other.

Revealed preferences. Slembeck and Tyran (2004) find that the ‘three-door anomaly’, a systematic violation of payoff-maximizing individual behavior in an experimental game, is eliminated when individuals have access to contextual cues in the form of communications from competing, better-performing participants in the game.

Stated preferences. Haward et al. (2008) presented subjects with the (fortunately) unfamiliar hypothetical choice between resuscitation and palliative (or “comfort”) care for extremely premature infants. However, even an impoverished description provided enough information for individuals who describe themselves as “highly religious” to resist a common framing manipulation. Namely, they responded consistently whether the outcomes were described in terms of survival and health or death and disability. Thus, some preferences reveal themselves, as soon as individuals are asked to state them.

4.4. Unfamiliar private goods, without informative contextual cues

Life sometimes surprises people with unfamiliar choices without any clarifying context. Experimental researchers do that as a matter of course, as they create novel formulations
designed to inform theoretical questions. Indeed, there can be a “curse of cleverness,” whereby the more innovative a task is, the less individuals can perform it consistently. When studies seek to highlight processes that are not easily seen in everyday life, the prevalence of (in)consistent preferences in them says little about their prevalence under more normal circumstances.

**Revealed preferences.** Preferences in this domain have been shown to be easily influenced by irrelevant information and not appropriately sensitive to relevant information. Behavioral economists have amply identified cases in which market choices are inconsistent with one another, not to mention consumers’ best interests (Ariely, 2008; Thaler and Sunstein, 2008). For example, Ariely et al. (2003) showed that willingness to pay for goods like bottles of wine (by students unfamiliar with the specific kinds) was affected by obviously arbitrary numbers (digits from their social security cards), to which their attention had previously been drawn. Tilman and Slembeck (2004) found that participants in an unfamiliar game did not appropriately update their preferences, given new information, unless they had access to informative contextual cues from better informed participants (cf. 4.3). One non-monetary example is the much greater rate of organ donation when it is the default, compared to where one must opt into it (Johnson et al., 2003). Thaler and Sunstein (2008) show the power of such defaults, with unfamiliar private goods and impoverished contexts.

**Stated preferences.** Studies with lightly described private goods are a staple of stated preference studies designed to demonstrate respondents’ sensitivity to normatively irrelevant cues. For example, Coupey et al. (1998) show sensitivity (preference reversals) to task details that should be irrelevant (choice vs. matching format) in stated preferences for unfamiliar products like “reverse osmosis filters.” Even when the good is ostensibly familiar (e.g., a coffee mug in an endowment effect study; Kahneman et al., 1990), the social setting of the specific experiment will not be. If the preference applies to an interpersonal exchange, then the good cannot be isolated from it – meaning the choice in an experiment is unfamiliar even when the good is common.

One operational test for whether cues are “informative” is whether preferences are affected by adding them in a non-coercive way. For example, Huffman et al. (2007) found that preferences for GM food stabilized when accompanied by information about the positions of politically involved third parties, who could be assumed to have worked through the issues from their perspective.

4.5. Familiar public goods, with informative contextual cues

At the end of idealized political processes, the public goods at issue are familiar and citizens’ choices are informed (e.g., Lupia and Matsusaka, 2004). That reality, of course, is always something less. Indeed, political scientists often focus on “issue publics,” individuals who care enough to familiarize themselves with a public good and its context (e.g., how will its costs and benefits be distributed, what precedents will it set). The issue public is held to interpret issues for others sharing their values. Whether even they can master the specific issue and context is an empirical question (Fischhoff, 1991).

**Revealed preferences.** Preferences for public goods are most clearly revealed in behaviours like voting, canvassing, and making political contributions. Experiments are uncommon, given that public goods often become familiar through political processes
that are not amenable to the experimental manipulations needed for consistency tests. As a result, such tests tend to be limited to choices about contributions to public goods. Such studies sometimes find inconsistent revealed preferences. For example, even after a withering national election campaign, voter turnout (a measure of caring about the result) could still be increased by several percentage points by the simple intervention of asking registered voters where they would be prior to voting (Nickerson and Rogers, 2010). Envisioning that place and the path from it to the polls provided context that was otherwise missing. A growing research area examines how alternative frames can affect charitable contributions for seemingly familiar public goods with ample contextual cues (e.g., people give more for identified beneficiaries than for the general causes that they represent; Small et al., 2007).

Stated preferences. Attitude surveys about public goods tend to provide few informative contextual cues, in deference to reducing respondents’ cognitive load. As a result, they will be discussed below (4.6). Chong and Druckman (2007) found that, even with relatively familiar public policy issues, which had been widely covered in the news media, stated preferences were still subject to framing effects, until respondents received (forged) arguments from political information providers (see also Druckman, 2004; Druckman and Nelson, 2003; Tomz and Sniderman, 2005).

Although rarely used in experimental economics or psychology, deliberative processes are often used in applied settings to help people to articulate their preferences, by hearing and reflecting on alternative perspectives, chosen to provide contextual information. Such clarifying interactions are central to decision analysis (Clemen & Reilly, 2002), which continues until clients’ preferences pass consistency tests. Morgan et al. (2001) describe the consistency achieved through a complex group deliberation about risk priorities. Well-designed deliberative processes show the limits to consistency, while providing only indirect evidence regarding its natural occurrence.

4.6. Familiar public goods, without informative contextual cues

Preferences here involve public goods that may or may not seem to require the additional information provided by contextual cues.

Revealed preferences. Political advertising counts on the malleability of voting behavior for issues that have been regularly, but only superficially aired. Its impacts are limited by the power of party identification. People do not need to know very much about specific issues, if they follow party leaders who take consistent positions. Political scientists study how individuals respond to such naturally occurring circumstances. Researchers are rarely in a position to create them.

Stated preferences. Public opinion polls normally fall into this category, letting the issues speak for themselves. Polls released to the general public sometimes elicit the strength of respondents’ preferences, but rarely examine their internal consistency. Proprietary polls done for political candidates often look for inconsistency, seeking the context that evokes the most desirable preferences. Several of the experimental studies on political behaviour mentioned in Section 4.5 found that, absent cues such as party labels, preferences tended to be inconsistent with one another (Tomz and Sniderman 2005) or with expressed political attitudes (e.g. Druckman and Nelson 2003).
4.7. *Unfamiliar public goods, with informative contextual cues*

It is a triumph of political deliberations when they take public goods from obscurity to fully informed status. As mentioned, the reality is often less satisfying. Even voters who know the main arguments of both sides still may not be comfortable making their decisions without considering contextual cues, such as party labels.

*Revealed preferences.* In actual voting, although experimentally manipulating information is impossible, it is possible to assess the consistency of preferences for voters with different amounts of contextual information. Lupia (1994) finds that otherwise uninformed voters who knew the positions of interested parties voted as consistently as well-informed voters with similar stakes riding on an issue.

*Stated preferences.* Researchers who present unfamiliar public good issues often provide little context, preferring to evoke intuitive responses. Indeed, many social scientists are averse to providing context, lest they unnaturally bias or deepen respondents’ preference (Fischhoff, 1991). Other researchers embrace *reactive measurement*, seeking to change respondents is a disciplined way. For example, deliberative polling (Fishkin, 1991; Luskin et al., 1997) tries to present balanced perspectives on an unfamiliar issue, in the context of group discussions moderated to facilitate airing and clarifying of alternative perspectives. Like decision analysis, these processes proceed until they reach the most consistent preferences possible, hence provide only indirect evidence regarding the natural prevalence of consistent preferences.

In a demonstration of the power of trusted authorities to provide useful contextual cues for unfamiliar public goods, Druckman (2001) found that forged party positions reduced framing effects with Tversky and Kahneman’s (1981) Asian disease problem. In experimental studies, providing actual party positions afforded hypothetical voters enough contextual information to form consistent preferences, as reflected in their sensitivity to price (Schläpfer et al., 2008). The predictive validity of many pre-election polls suggests that they can evoke the context of the actual voting, although they seldom assess how well voters understand the context (Vossler and Kerkvliet 2003).

4.8 *Unfamiliar public goods, without informative contextual cues*

Psychologists have long known that most survey respondents offer some opinion on almost any question, no matter how unfamiliar the topic or how impoverished the context cues. A classic example is that 70% of survey respondents expressed an opinion on a fictitious Metallic Metals Act although no such act existed and, therefore, the respondents could have no actual knowledge (Gill, 1947; cited in Plous, 1991, p. 55). Some such respondents do not want to admit ignorance. Some assume that they can guess what the issue is, based on the brief description.

In such cases, the good is not only unfamiliar, but undefined. Fischhoff and Furby (1988) present one framework for specifying choices well enough that consistent preferences are possible. If everything needs to be spelled out, then the cognitive load could be unmanageable. However, informative contextual cues can reduce that load by directing respondents to critical issues and proposing possible decision rules.

*Revealed preferences.* Despite theoretical arguments regarding the informational value of advertising, those using it often seek to manipulate contextual cues in their favor.
Their opportunities are greatest when the good is unfamiliar or misunderstood. When marketers are successful, the contextual cues determine which features are essential to the good’s definition. For instance, Great Britain’s close national referendum on joining the EU was reportedly tipped by framing it as “stay in Europe,” rather than “join Europe,” a distinction discovered through research into what is now called the “status quo bias.”

Contingent valuation surveys are sometimes described as producing results that will affect policies regarding public goods. If that claim is accepted, then responses might reveal (rather than just state) preferences. It also heightens theoretical concern about strategic responses, understating or overstating a good’s value, based on how respondents assess the opportunities for slanting results to their advantage (Gibbard, 1973; Satterthwaite, 1975; see also Green et al., 1998; Champ et al. 2002; Flores and Strong, 2007; Schläpfer and Bräuer, 2007; Schläpfer and Schmitt, 2007).

Stated preferences. Manipulation checks ask respondents to recall and make inferences about details of questions that they have just answered. Without them, there is no direct way of telling how familiar the issue was. The power of framing, and other context effects, arises from respondents’ need to rely on contextual cues, in order to articulate reasoned answers to incompletely specified questions. In many experimental studies, those cues are deliberately uninformative, so that researchers can demonstrate an “effect,” whereby people respond to irrelevant information. The psychological and contingent valuation literatures have many such examples (Tversky and Kahneman, 1981; Fischhoff and Manski, 1999; Lichtenstein and Slovic, 2005; McFadden 2001). These demonstrations show how inconsistent preferences can be with goods that are both incompletely specified goods and without informative contextual cues. Ironically, the same lack of information and context can also encourage insensitivity to relevant information, such as the quantity or ‘scope’ of a good (Boyle et al., 1994; McFadden, 1994). One of the few stated preference studies with manipulation checks found that preferences were more consistent when evaluated in terms of what respondents believed about the goods, compared to what was explicitly stated (Fischhoff et al., 1993).

4.9. Discussion

Thus, both consistent and inconsistent preferences have been found with both revealed and stated preferences, in most of these eight domains. As the discussion suggested, some domains are much more heavily studied than others, reflecting research opportunities (e.g., for experimental manipulations) and researcher predilections (e.g., for minimal or hypothetical descriptions). Because these predilections differ by academic discipline, researchers typically work in a subset of possible domains and see the patterns of consistency that it typically produces. It would only be natural, if they came to see such results as typical in the world.

The present analysis finds that preference consistency seems to increase as choices become better understood, whether through familiarity with the good or receipt of informative contextual cues. That is seen in contrasts across domains, created to vary on these two features, and within domains, for choices varying on these two factors. The analysis did not suggest any systematic impacts of the other two features: whether the good was public or private and whether preferences were revealed or stated. In studies (and in the world), public and private goods tend to vary in how well they are understood.
So do revealed and stated preference tasks. As a result, differences in preference consistency reflect those covariates (familiarity, informative contextual cues), not the public/private or revealed/stated distinction.

The following section examines whether these patterns emerge in the universe of preferences represented in one set of published studies.

5. Explaining “hypothetical bias” in stated preferences: a meta-analysis regression

The preceding analysis suggested that preference consistency depends on how well choice tasks are understood. That depends, in turn, on the familiarity of the good and the informativeness of the accompanying contextual cues. Preference consistency does not, however, depend on whether the good is private or public good or whether the preferences are revealed or stated.

The studies used as examples were selected to illustrate these cases. This section examines these patterns in a set of studies collected for other purposes. Specifically, it asks how well three of these features (public/private, familiarity, cue informativeness) predict a form of consistency defined by the fourth: whether stated and revealed preferences agree. A standard measure of such consistency is hypothetical bias, the ratio of willingness to pay expressed in stated preference tasks (with hypothetical choices) and in revealed preference tasks (involving actual payments). These ratios vary widely enough, across studies, to justify the search for covariates.

The set of studies is drawn from Murphy et al. (2005). They used OLS to predict hypothetical bias from six factors: whether the good is public or private, whether the preferences are from students or other (perhaps more strongly motivated) individuals, whether preferences are elicited from individuals or (possibly more thoughtful) groups, whether preferences were expressed or an open-ended scale or choices among fixed options that might suggest expected answers (e.g., possible prices); whether the design was between-subjects or within-subjects (which might have suggested expected answers); and whether preferences were somehow “calibrated” in order to be more consistent (see Murphy et al. 2005, p. 318). They considered 28 studies, with 83 observations, selected because of their accessibility in the peer-reviewed literature (see Appendix, Table A1).

Model 1a in Table 2 shows regression results for all 83 observations included in the original analysis. The only variable that strongly predicted a smaller hypothetical bias was using a within-subject design. Murphy et al. note that “attempts to identify other factors that may be associated with hypothetical bias yielded mixed results.”

For these purposes, two of these six predictors seem problematic. As mentioned, within-subject designs can signal the features that interest the researchers, thereby inducing consistency (Fischhoff and Bar Hillel, 1984). Calibration of responses attempts to eliminate the phenomenon of interest. As a result, we reran the regression after

---

1 The original dataset is available at: http://faculty.cbpp.uaa.alaska.edu/jmurphy/meta/meta.html#data (accessed 8 June, 2010.)
eliminating the 44 observations having within-subject designs or calibration. Model 2a shows that hypothetical bias is somewhat smaller with private goods and choice tasks.

We now repeat these regressions, adding the two variables central to our analysis: familiarity with the good and access to informative contextual cues. Table A in the Appendix shows our coding of the 28 studies (with 83 observations) on these two additional variables. Among them, 13 had unfamiliar goods without informative context, 7 had unfamiliar goods with informative context, 5 had familiar goods without informative context, and 3 had familiar goods with informative context.

Because Model 1b includes with-subject designs and calibration, we focus on Model 2b. In it, hypothetical bias is significantly smaller for familiar goods and for informative contexts. The significant interaction indicates that familiarity and context are somewhat redundant contributors to more consistent preferences: contextual cues reduce hypothetical bias less with familiar goods than with unfamiliar ones. Conversely, familiarity reduces bias less when people have access to contextual cues.

[Table 2 about here]

6. Conclusions

Economists have traditionally made a strong distinction between revealed and stated preferences. The former are usually held to meet the axiomatic assumptions for preferences, while the latter are often dismissed for their perceived inconsistency. However, as seen in Section 4, both consistent and inconsistent choices are found with both revealed and stated preferences. We propose that consistent preferences are more likely when individuals understand tasks better, regardless of whether the choices involve public or private goods and regardless of whether the preferences are revealed or stated.

We propose further that understanding can come two complementary routes. One is familiarity with the good. The second is receipt of informative contextual cues. If revealed preferences studies are more likely to demonstrate consistent preferences, it is because those studies are more likely to involve well-understood tasks. The analysis in Section 4 provides reasons why this might be the case. For example, stated preference studies can (and often do) use choices invented for research purposes. Conversely, revealed preference tasks do sometimes occur in life, with relatively fixed and familiar properties and context.

We supplement this illustrative discussion with statistical analysis of a sample of studies, collected by Murphy et al. (2005) for other purposes (thereby reducing the chances that they were selected to confirm our hypotheses). In their full sample, the only predictor of choice consistency was using a within-subject design, which can signal how to be consistent. Removing those studies, along with ones that “calibrated” responses to be more consistent, revealed two predictors of consistency: using private (rather than public) goods and eliciting preferences with choice (rather than rating scale) responses. However, those predictions vanished when familiarity and contextual cues were added to

2 The dropped observations include 17 within-subject observations from Sinden (1988), multiple observations in Blumenschein et al. (2001) and Johannesson et al. (1998), and calibrated observations by List (2001) and List (2003), cited in Murphy et al. 2005.

3 The correlation between the two variables is r=0.16.
the equations. Thus, the other two predictors seem to have mattered because they were more likely in studies postulated as having better understood tasks.

If this result is robust, then concern over revealed versus stated preferences has been misplaced, obscuring the more important question of when people understand tasks well enough to make consistent responses. We propose that one class of determinants has to do with how familiar individuals are with a good. Whether a good is public or private may predict familiarity, for researchers who focus on familiar private goods. Private goods may also have a more natural set of informative contextual cues, such as prices on competitive markets or popularity among personally relevant individuals. However, those correlations are not exclusive. Political campaigns and policy debates can make public goods quite familiar, as well as generate party positions providing important contextual cues. Similarly, researchers and marketers can create entirely novel private goods for decision makers’ consideration.

From this perspective, the critical research question becomes what are the conditions that favor understanding choice tasks and, hence, consistent preferences. The answers to that question can be used prescriptively, when the goal is to create those conditions (Fischhoff et al. 1999; Keeney and Raiffa, 1993; Payne et al. 1999). The answers can be used descriptively when the goal is to understand how well natural circumstances have allowed individuals to master choice tasks (Fischhoff et al., 1980; Lichtenstein and Slovic, 2005). In a mature research program, the two activities would be integrated. Descriptive research would direct prescriptive efforts to make tasks clear, then assess their success. Prescriptive work would identify barriers to understanding that could benefit from and inform descriptive research. That integrated research program could, for example, examine the role of consequentiality of choices on their consistency. In other domains, it has proven theoretically productive to ask when increased incentives enhance or degrade performance or have no effect at all, absent help in understanding tasks (Camerer and Hogarth, 1999; Lerner and Tetlock, 1999.; Milkman et al., 2009).

The integration of descriptive and prescriptive research has long been part of the behavioural decision research, which grew out of mathematical, cognitive, and engineering psychology in the 1950s (Edwards, 1954; Kahneman et al., 1982 vonWinterfeldt and Edwards, 1986). It seems to be a historical coincidence that such integration has lagged as economics has strengthened its behavioural foundations. One barrier has been the theoretically founded scepticism regarding stated preferences, reinforced by the fact that so many stated preference studies have involved unfamiliar goods presented with limited contextual cues (Manski 2000). A second barrier has been the focus of revealed preference studies involving familiar goods where these issues were less relevant. A third barrier has been the institutional pressures to defend many stated preference estimates, for example, when contingent valuations are used to defend damage settlements. Such settings encourage doing the best study possible, then denying any residual problems.

All researchers must work within the constraints of their funding and their discipline’s norms. However, all those concerned about assessing and informing preferences could benefit from better fundamental understanding of the conditions that favor consistent preferences. That basic science can exploit and inform research in economics, psychology, political science, survey research, and other fields. It raises questions such as what experiences make a good familiar enough that people can
anticipate their future experiences, which sources are trusted for providing contextual cues, which cues are most informative, and what insight do individuals have into how well they have articulated their own preferences.
References


Nickerson D. W. and Rogers, T., 2010. Do you have a voting plan?: implementation intentions, voter turnout, and organic plan making. Psychological Science, 21 (2), 194-9


Table 1. Classification of choice domains.

<table>
<thead>
<tr>
<th>Private vs. public good</th>
<th>Familiarity (own experience)</th>
<th>Informative contextual cues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td>familiar</td>
<td>present</td>
</tr>
<tr>
<td></td>
<td></td>
<td>absent</td>
</tr>
<tr>
<td></td>
<td>unfamiliar</td>
<td>present</td>
</tr>
<tr>
<td></td>
<td></td>
<td>absent</td>
</tr>
<tr>
<td>Public</td>
<td>familiar</td>
<td>present</td>
</tr>
<tr>
<td></td>
<td></td>
<td>absent</td>
</tr>
<tr>
<td></td>
<td>unfamiliar</td>
<td>present</td>
</tr>
<tr>
<td></td>
<td></td>
<td>absent</td>
</tr>
</tbody>
</table>
Table 2. Ordinary least squares estimates of models explaining the ratio of hypothetical to actual willingness to pay.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full dataset</th>
<th></th>
<th></th>
<th>Trimmed dataset&lt;sup&gt;a&lt;/sup&gt;</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1a</td>
<td>Model 1b</td>
<td>Model 2a</td>
<td>Model 2b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
<td>t-ratio</td>
<td>Coeff.</td>
</tr>
<tr>
<td>Constant</td>
<td>3.628***</td>
<td>3.978</td>
<td>2.733***</td>
<td>2.700</td>
<td>5.628***</td>
<td>3.617</td>
</tr>
<tr>
<td>Student</td>
<td>0.907</td>
<td>0.724</td>
<td>-2.434</td>
<td>-1.635</td>
<td>2.221</td>
<td>1.030</td>
</tr>
<tr>
<td>Group</td>
<td>0.955</td>
<td>0.771</td>
<td>4.906***</td>
<td>2.834</td>
<td>1.551</td>
<td>0.812</td>
</tr>
<tr>
<td>Private</td>
<td>-0.065</td>
<td>-0.082</td>
<td>3.390***</td>
<td>3.193</td>
<td>-2.935*</td>
<td>-1.935</td>
</tr>
<tr>
<td>Choice</td>
<td>-1.167</td>
<td>-1.390</td>
<td>0.165</td>
<td>0.192</td>
<td>-3.134**</td>
<td>-2.066</td>
</tr>
<tr>
<td>Calibrate</td>
<td>-1.838*</td>
<td>-1.846</td>
<td>-2.466***</td>
<td>-2.696</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cues (C)</td>
<td>-1.991*</td>
<td>-1.704</td>
<td>-3.665*</td>
<td>-1.827</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar (F)</td>
<td>-8.030***</td>
<td>-4.341</td>
<td>-9.062***</td>
<td>-3.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C x F</td>
<td>5.028**</td>
<td>2.402</td>
<td>6.723**</td>
<td>2.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>83</td>
<td>83</td>
<td>39</td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.08</td>
<td>0.26</td>
<td>0.13</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Within-subject comparisons and observations with “calibrated” values excluded.

Notes: Significance levels: *: $p<0.1$; **: $p<0.05$; ***: $p<0.01$. 
## Appendix Table A. Description of goods and choice contexts used in the study sample of the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Good and choice context</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blumenschein, et al. (1997)</td>
<td>lab sunglasses to economics students</td>
<td>0</td>
</tr>
<tr>
<td>Blumenschein et al. (2001)</td>
<td>asthma management program to patients in “scientific study”, general population</td>
<td>0</td>
</tr>
<tr>
<td>Bohm (1972)</td>
<td>preview of TV program at broadcast station, general population</td>
<td>1</td>
</tr>
<tr>
<td>Botelho and Costa Pinto (2002)</td>
<td>contribution to NGO information campaign on river otters, at university, students</td>
<td>0</td>
</tr>
<tr>
<td>Boyce et al. (1989)</td>
<td>common house plant in experiment, university staff at university</td>
<td>1</td>
</tr>
<tr>
<td>Brown et al. (1996)</td>
<td>contribution to remove roads along Grand Canyon rim, mail survey/collection by university</td>
<td>0</td>
</tr>
<tr>
<td>Brown and Taylor (2000)</td>
<td>contribution to NGO environmental program, at university</td>
<td>0</td>
</tr>
<tr>
<td>Cameron, et al. (2002)</td>
<td>green electricity contribution, in established renewable energy and tree planting program (fixed 6$ surcharge)</td>
<td>0</td>
</tr>
<tr>
<td>Carlsson and Martinsson (2001)</td>
<td>contribution to 3 environmental projects “rain forest”, “Mediterranean”, “Baltic sea”, students prior to lecture</td>
<td>0</td>
</tr>
<tr>
<td>Champ et al. (1997)</td>
<td>contribution to remove roads along Grand Canyon rim, mail survey/collection by university</td>
<td>0</td>
</tr>
<tr>
<td>Champ and Bishop (2001)</td>
<td>green electricity donation, actual wind power program of electric company, fixed surcharge</td>
<td>1</td>
</tr>
<tr>
<td>Duffield and Patterson (1992)</td>
<td>contribution to Montana river mgt., contingent valuation survey context</td>
<td>0</td>
</tr>
<tr>
<td>Frykblom (1997)</td>
<td>Swedish national atlas, volume “the environment”, prior to lecture</td>
<td>1</td>
</tr>
<tr>
<td>Frykblom (2000)</td>
<td>Swedish national atlas, volume “the environment”, prior to lecture</td>
<td>1</td>
</tr>
<tr>
<td>Heberlein and Bishop (1986)</td>
<td>hunting permits to hunters, issued by Department of Natural Resources</td>
<td>1</td>
</tr>
<tr>
<td>Johannesson et al. (1998)</td>
<td>box of chocolates, displayed in front of students</td>
<td>1</td>
</tr>
<tr>
<td>Kealy et al. (1988)</td>
<td>familiar candy bars (Cadbury)</td>
<td>1</td>
</tr>
<tr>
<td>List (2001)</td>
<td>sports cards in existing market (list price of $200-250)</td>
<td>0</td>
</tr>
<tr>
<td>List (2003)</td>
<td>sports cards in existing market (book value of $12)</td>
<td>0</td>
</tr>
<tr>
<td>List and Shogren (1998)</td>
<td>sports cards in existing market ($350)</td>
<td>0</td>
</tr>
<tr>
<td>Loomis et al. (1997)</td>
<td>wildlife art print displayed to subjects, clerical and administrative staff on campus, auction context</td>
<td>1</td>
</tr>
<tr>
<td>MacMillan, et al. (1999)</td>
<td>contribution to nature reserve, mail survey experiment, although mimicking regular collection</td>
<td>0</td>
</tr>
<tr>
<td>Murphy et al. (2002)</td>
<td>contribution to environmental NGO, students at university</td>
<td>0</td>
</tr>
<tr>
<td>Neill et al. (1994)</td>
<td>painting (list price $75), historical map replica (list price $20) shown to subjects, students at university</td>
<td>0</td>
</tr>
<tr>
<td>Sinden (1988)</td>
<td>contribution to soil conservation program, students in class</td>
<td>0</td>
</tr>
<tr>
<td>Spencer et al. (1998)</td>
<td>contribution to pond water monitoring program, student respondents</td>
<td>0</td>
</tr>
<tr>
<td>Vossler et al. (2003)</td>
<td>pre-election poll on collectively provided public good, issue subject to public debate, general population</td>
<td>0</td>
</tr>
<tr>
<td>Vossler and Kerkvliet (2003)</td>
<td>pre-election poll on collectively provided public good, issue subject to public debate, general population</td>
<td>0</td>
</tr>
<tr>
<td>Working Papers of the Socioeconomic Institute at the University of Zurich</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1007</strong> When Are Preferences Consistent? The Effects of Task Familiarity and Contextual Cues on Revealed and Stated Preferences, Felix Schläpfer, Baruch Fischhoff, August 2010, 20 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1006</strong> Golden Balls: A Prisoner’s Dilemma Experiment, Donja Darai, Silvia Grätz, July 2010, 47 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1005</strong> Probability Weighting as Evolutionary Second-best, Florian Herold, Nick Netzer, July 2010, 32 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1004</strong> Trade Openness, Gains from Variety and Government Spending, Sandra Hanslin, April 2010, 46 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1002</strong> Preferences for Health Insurance in Germany and the Netherlands – A Tale of Two Countries, Peter Zweifel, Karolin Leukert, Stephanie Berner, March 2010, 22 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1001</strong> Convex Treatment Response and Treatment Selection, Stefan Boes, January 2010, 28 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0917</strong> How much do journal titles tell us about the academic interest and relevance of economic research? An empirical analysis. Felix Schläpfer, December 2009, 14 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0916</strong> Fine Tuning of Health Insurance Regulation: Unhealthy Consequences for an Individual Insurer, Johannes Schoder, Michèle Sennhauser, Peter Zweifel, August 2009, 18 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0915</strong> Capping Risk Adjustment?, Patrick Eugster, Michèle Sennhauser, Peter Zweifel, September 2009, 27 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0913</strong> Copula-based bivariate binary response models, Rainer Winkelmann, August 2009, 26 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0912</strong> Simulating WTP Values from Random-Coefficient Models, Maurus Rischatsch, July 2009, 6 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0911</strong> Physician dispensing and the choice between generic and brand-name drugs — Do margins affect choice?, Maurus Rischatsch, Maria Trottmann, July 2009, 15 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0909</strong> Social Mobility and Preferences for Income Redistribution: Evidence from a Discrete Choice Experiment, Ilja Neustadt, Peter Zweifel, July 2009, 31 p.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0907</strong> Competitive Screening in Insurance Markets with Endogenous Wealth Heterogeneity, Nick Netzer, Florian Scheuer, April 2009, 28 p.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
0905 Patents versus Subsidies – A Laboratory Experiment, Donja Darai, Jens Großer, Nadja Trhal, March 2009, 59 p.
0904 Simple tests for exogeneity of a binary explanatory variable in count data regression models, Kevin E. Staub, February 2009, 30 p.
0815 Delay and Deservingness after Winning the Lottery, Andrew J. Oswald, Rainer Winkelmann, December 2008, 29 p.
0810 Why does the amount of income redistribution differ between United States and Europe? The Janus face of Switzerland; Sule Akkoyunlu, Ilja Neustadt, Peter Zweifel, September 2008, 32 p.
0718 Monoplistic Screening under Learning by Doing, Dennis Gärtner, December 2007, 29 p.