STATED PREFERENCES FOR PUBLIC SERVICES: A CLASSIFICATION AND SURVEY OF APPROACHES

Felix Schläpfer*

Department of Economics and Management, Kalaidos University of Applied Sciences

Abstract. The wider range of stated preference approaches to value public goods has not been systematically reviewed in recent years. The objective of this paper is to provide an overview of this literature and to evaluate the strengths and limitations of alternative approaches. Since the public referendum has served as a ‘blueprint’ for survey design, two key dimensions by which many surveys differ from the public referendum are used for a simple classification of approaches. This yields eleven approaches, including different variants of micro-based demand surveys, referendum surveys, budget allocation surveys and contingent valuation surveys. Their evaluation in terms of the preference information they produce and the assumptions they require suggests there is no single preferred approach. Instead, each approach has its characteristic profile of strengths and limitations which follow from how it strikes the balance between the conflicting goals of measuring entire willingness-to-pay distributions and presenting manageable, credible and incentive compatible questions. Ultimately, judgments about the suitability of alternative approaches for specific objectives should rely on empirical evidence. Progress in the field could greatly benefit from a routine implementation of powerful experimental validity tests in applied work.

Keywords. Budget allocation; Contingent valuation; Demand; Public goods; Referendum; Stated preferences

1. Introduction

Individual preferences for public services are central to the economic analysis of decisions in the public sector, but their measurement remains challenging. Preferences for public goods and services cannot be directly observed on markets. In the political domain, a fundamental institution in which individual preferences for public services are expressed and aggregated is voting on public referenda (Bowen, 1943; Bergstrom et al., 1982). However, in many settings and for many services of interest, voting behaviour is usually not available for analysis. In such cases, researchers are increasingly using surveys of stated preferences to measure individual values of public goods and services (Mitchell and Carson, 1989; Bateman et al., 2002; Carson, 2012; Kling et al., 2012).

Traditionally, the term ‘stated preferences’ has been used to describe preferences obtained from surveys as opposed to ‘revealed preferences’ which are derived from choices that are directly linked to transactions of goods or services in markets or in the political domain. Interestingly, however, in the recent literature on the valuation of public goods and services, the term ‘stated preferences’ is used almost synonymously with one specific approach – the contingent valuation (CV) approach, including a variant called choice

*Corresponding author contact email: felix.schlaepfer@kalaidos-fh.ch; Tel: +41 44 200 19 02.
For instance, in their paper ‘A common nomenclature of stated preference elicitation approaches’, Carson and Louviere (2011) write (p. 544): ‘Contingent Valuation can be thought of as short-hand for using SP questions to address standard applied microeconomic welfare issues characterizing public goods (pure and quasi-public) in arenas like culture, environment and health’. Accordingly, the authors consider only variants of the contingent valuation approach in their classification of stated preference approaches. Likewise, the monographs by Mitchell and Carson (1989) and Bateman et al. (2002) consider only contingent valuation. Kling et al. (2012, p. 8) explicitly state that they use contingent valuation (including choice experiments) and stated preferences interchangeably.

While contingent valuation is now a dominant approach, surveys of stated preferences are more diverse than the recent literature suggests. Until the 1970ies and 1980ies, the contingent valuation approach was only one among a variety of survey-based approaches in applied welfare analysis. These approaches include various forms of budget allocation surveys (e.g. Pendse and Wyckoff, 1976) and demand surveys (e.g. Bergstrom et al., 1982) which differ from contingent valuation in fundamental ways. More recently, the particular types of preference questions and information provision used in contingent valuation have been questioned based on evidence of anchoring effects and other choice inconsistencies (e.g. Fischhoff, 1991; McFadden, 1994; Kahneman et al., 1999; Ariely et al., 2003) and potential lack of scenario credibility (Champ et al., 2002, Flores and Strong, 2007). Based on these findings, new types of preference questions and new ways of information provision have been proposed (e.g. Schlapfer and Schmitt, 2007; Getzner, 2012).

Considering a broader range of approaches in stated preference research, including survey approaches that are rarely used today, may be of interest for several reasons:

1. In spite of large efforts in research over the past decades, the currently dominant approach in stated preference research – asking discrete choice willingness-to-pay (WTP) questions with counterfactual cost figures in self-contained surveys – remains contentious within and outside the economics discipline (e.g. Ariely et al., 2003; Diamond and Hausman, 1994; McFadden, 1999; Hausman, 2012; Kling et al., 2012).

2. Too much reliance on one contentious approach may prevent stated preference research from developing its full applied potential in fields where it is still rarely used (Sunstein, 2014). Policy makers would likely be interested in additional types of stated preference information if they knew more about the alternative approaches which are largely neglected in courses, textbooks and reports (USEPA, 2009).

3. There exists virtually no literature investigating or only discussing the relative strengths of contingent valuation and alternative approaches in the analysis of preferences for public goods. This may be partly due to the fact that the approaches have developed in separate strands of the literature (see e.g. Brookshire and Crocker, 1981).²

4. The state of affairs in contingent valuation is unsatisfactory in light of recent challenges from behavioural economics (e.g. Ariely et al., 2003; Sunstein and Thaler, 2003). These challenges are now routinely acknowledged (Loomis, 2011; Carlsson, 2012) without, however, being substantially addressed (e.g. McFadden, 1999; Alevy et al., 2011).

The objective of this paper is thus to provide an overview of the larger set of alternatives in stated preference research, to start identifying and discussing the specific strength and limitations of the different approaches for a range of applied purposes, and to offer conclusions for future research.

Following a simple classification of stated preference approaches in Section 2, the subsequent sections address five specific questions: which types of preference information do the different approaches produce?; which specific challenges do they address?; which assumptions do they make?; which are specific strength and limitations of the approaches; and which approaches may be most useful for particular applied objectives? A final section offers discussion and conclusions.
Table 1. Classification of Survey Approaches.

| Price information | Self-contained survey ('homegrown preferences') | Survey with information on the political support (IPS) of proposals |
|-------------------|-----------------------------------------------|---------------------------------------------------------------|
| Actual price      | -Micro-based demand survey                     | -Micro-based demand survey with IPS                          |
|                   | -Budget allocation survey                      | -Budget allocation survey with IPS                           |
|                   | -Referendum survey                             | -Referendum survey with IPS                                   |
| Randomly assigned but realistic price | -Budget allocation survey with experimental variation of percentage or per-unit charges | -Budget allocation survey with experimental variation of percentage or per-unit charges, with IPS |
|                   | -CV survey with experimental variation of percentage or per-unit charges | -CV survey with experimental variation of percentage or per-unit charges, with IPS |
| Randomly assigned and unrealistic price | -Standard CV survey with experimental variation of individual costs | -Not feasible since providing IPS on counterfactual proposals is not feasible |

2. A Simple Classification of Stated Preference Survey Approaches

2.1 Criteria for Classification

In the contingent valuation literature the public referendum has been recommended as a ‘blueprint’ for survey design (Arrow et al., 1993). Following this recommendation, CV surveys today typically use binary questions – the so called ‘referendum format’, and they specify the proposed policies in much detail, including a realistic scenario on how the policy will be financed (Kling et al., 2012). However, a contingent valuation survey with binary referendum questions is not simply the sample-survey analogue of an (advisory) public referendum. Standard contingent valuation surveys differ from the public referendum in at least two important additional ways: (1) The specified price of the public service is hypothetical in the sense that it is not the actual price that an individual would pay if the proposed policy is implemented but a randomly assigned amount as required for the statistical analysis of WTP (Hanemann, 1984). (2) The survey process is self-contained; while decision makers in referenda about unfamiliar public issues typically draw on information about the issue positions of opinion leaders, political parties or interest groups, the respondents in CV surveys do not have comparable information at hand (e.g. Shapiro and Deacon, 1996).

If the goal of stated preference surveys is essentially to sample voter preferences (cf. Arrow et al., 1993; Hanemann, 1994), then the two dimensions along which stated preference surveys may differ from the ‘blueprint’ deserve close attention. In the following, these dimensions – actual versus randomly assigned prices and presence versus absence of information about the distribution of political support of the proposed policies – are used for a simple classification of existing stated preference approaches (Table 1). Stated preference approaches here includes representative sample surveys in which individuals make choices between (mainly) public goods or services and money and which are used to determine, among a population, preference parameters that are relevant to the provision of public goods or services such as the mean, median or upper/lower bounds on WTP, preferred quantities, demand elasticities, or pluralities (within groups) supporting a specified policy. This definition is wide but excludes for instance
happiness surveys (e.g., Levinson, 2012) and deliberative monetary valuation approaches which can hardly be implemented with large representative samples (e.g., Lo and Spash, 2013).

2.2 Prices in Binary Preference Questions: Actual Versus Randomly Assigned

The voters in public referendum decisions are confronted with actual aggregate project costs. For instance, if my city council proposes to spend an annual amount equivalent to 0.5% of the city budget to purchase land development rights, then I anticipate I would have to contribute with an amount equivalent to roughly 0.5% of my annual city tax. In contrast, prices of public services in dichotomous preference questions or in questions about preferred budget allocations may be specified in alternative ways. The first option – similar to the public referendum – is to specify, to the best of the researcher's knowledge, the true price or at least a realistic estimate of what the respondent would have to pay for the proposed policy (e.g., Vossler et al., 2003). A second option is to randomly assign different prices to different respondents following experimental design techniques in order to be able to statistically identify WTP distributions (Hanemann, 1984). A third, intermediate approach is to specify the price as a percentage of current taxes or as a price per unit of service which is drawn from an experimental design of realistic percentages or per-unit prices to statistically identify WTP. To illustrate the intermediate option, the price would be specified as a (randomly assigned) percentage change in a tax which might be drawn from a vector (0.2%, 0.4%, 0.6%, 0.8%, 1%). This percentage variation of the price would correspond with a wide range of relevant individual amounts: For individuals with a current tax bill of $1000, the price range would be between $2 and $10. For individuals with a tax bill of $10,000, the price range would be between $20 and $100. Hence, specification of percentage or per-unit prices enables the researcher to examine a large range of 'bids' while maintaining realism of individual amounts. The implication of this intermediate specification is that there is no contradiction between the specified payment vehicle (such as an income tax) and the randomly assigned prices (Schläpfer and Schmitt, 2007).

The nature of the price information in surveys – randomly assigned versus actual – relates to important conceptual alternatives in preference elicitation. A random assignment of prices is required if the aim is not only to identify the proportion of respondents approving a proposed service but also to statistically identify the distribution of WTP for that service (Hanemann, 1984). Specifying the true or 'actual' price in a dichotomous question is required if the analyst takes issues of scenario credibility and strategic answering into account. With randomly assigned prices, the respondents may not believe that the specified prices are the true amounts they would have to pay if the policy is implemented (Champ et al., 2002). Such respondents may then 'update' the price and answer different questions than those intended by the researcher (Flores and Strong, 2007). They may even answer strategically, since dichotomous preference questions are incentive compatible only if the alternatives (including their prices) are the true alternatives (Gibbard, 1973; Satterthwaite, 1975), or if the respondents at least believe that they are (Green et al., 1998; Carson and Groves, 2007). That bid credibility is a relevant issue is shown in a study by Champ et al. (2002) where 42% of the respondents did not share that belief.

2.3 Access to Information about the Political Support for the Proposed Service

In public referendum decisions about public services, individuals may draw on two basic types of information: One is technical information about the attributes or characteristics of the alternatives such as price, quality, duration etc. The other is information about who is in favor of implementing the proposed policy – e.g. which individuals, political parties and other interest groups. Accordingly, stated preference surveys may differ in terms of the types of information that are available to the respondents. One option is to provide only factual information about the alternatives. This is the situation when respondents are asked about their preferences for a novel policy that has not been publicly discussed. The policy is described in
a ‘self-contained’ survey (e.g. Arrow et al., 1993) and the respondents use that information to formulate their ‘homegrown’ preferences (e.g. Cummings et al., 1995). Another option is to allow for both factual information and information about the distribution of political support of the policy. This is possible if the proposed policy has been publicly discussed (e.g. Vosser et al., 2003). In many cases, of course, the researchers are precisely interested into the preferences for specific policy proposals that have never been publicly discussed. In such cases, information about the distribution of political support may be collected by the researcher and provided as supplementary information in the survey instrument (e.g. Schläpfer and Schmitt, 2007; Getzner, 2012).

Stated preference questions in the presence vs. absence of contextual political advice require different assumptions on individual rationality (Smith, 2003; Sunstein and Thaler, 2003). Behavioral research by political scientists demonstrates that ‘cues’ containing information about the distribution of political support for proposed policies can support successful heuristic decision making (e.g. Lupia and McCubbins, 1998; Lupia and Matsusaka, 2004). When such cues are available, the effects of arbitrary framing appear to be much decreased or eliminated (Lupia, 1994; Druckman, 2001; Schläpfer, 2011; Bechtel et al., 2015). Surveys that do not provide information about the distribution of political support are implicitly based on the assumption that respondents do not need such information to articulate stable preferences for complex public goods. A large body of behavioral research by psychologists, economists and political scientists now suggests that assumption is no longer tenable (Tversky and Kahneman, 1981; McFadden, 1994; Ariely et al., 2003). Of course, if effects of arbitrary framing such as anchoring effects can be decreased or eliminated by political advice or ‘cues’, this raises the question whether those cues introduce a new source of bias, and how large that bias might be relative to the bias due to arbitrary anchors (see Section 6.7).

3. Stated Preference Approaches and the Types of Preference Information they Produce

3.1 Micro-Based Demand Surveys

In micro-based demand surveys for public services respondents are asked only about their preferred direction of a change in specific public services. Individuals are thus asked whether they would prefer ‘more’, ‘less’ or ‘the same amount’ of a public service given that tax payments would have to be adjusted accordingly. Given responses from different local jurisdictions with different current expenditures and further survey-based information on individual incomes, tax prices and further socioeconomic information it is possible to derive demand elasticities for public goods. The analytical framework to analyze this type of preference information was developed by Bergstrom et al. (1982) with the specific aim to estimate demand elasticities in cases where voting data are unavailable for that purpose and where respondents may find it difficult to answer quantitative questions about the preferred amounts of public services. Further applications in the US include Gramlich and Rubinfeld (1982), Rubinfeld et al. (1987) and Rubinfeld and Shapiro (1989). Bergstrom et al. (1988) used the same data to estimate the marginal rates of substitution for public goods. Outside the US, Preston and Ridge (1995) use the same type of survey data (from the British Social Attitudes Survey) to estimate demand elasticities for total public spending on local services. Rongen (1995) applied the approach to Norwegian survey data and Ahlin and Johansson (2001) to six local public services is Sweden.

3.2 Referendum Surveys

Referendum surveys ask binary (yes/no) questions about whether specific policy projects should go forward. This can be advisory referenda or simply opinion polls on tax expenditures (or ‘millage’) for proposed public services. Individuals are asked whether they approve or disapprove of a proposed public service, considering both their perceived tax increase and benefits from the public service.
The costs of the programme to the individual respondent are not necessarily specified. Typically, the context implies that the usual tax structure and sources of public revenue apply. If a proposed policy alternative is the true and only alternative policy that is available, the referendum question has desirable incentive properties (Gibbard, 1973; Satterthwaite, 1975). The analytical opportunities are the same as in the analysis of actual voting decisions using micro data (e.g. Rubinfeld, 1977; Fischel, 1979). What can be derived from the responses is the frequency distribution of individual WTP values that are above/below the perceived individual costs of the policy. Hence, for a given distribution of the (tax) costs, it is possible to derive the plurality of the population approving the policy. The sum of the implicit WTP values of the approving respondents can serve as a conservative lower bound estimate of the aggregate WTP for the policy.3 Interestingly, while many economic studies examine preferences in voting decisions, there are very few economic studies analyzing responses to simple millage referendum questions in surveys. An example is Vossler et al. (2003) who conducted such a survey with the aim of comparing the responses to the binary valuation questions with those in a subsequent actual voting decision.

If the survey questions focus on a single proposal (expenditure level), then it is only possible to estimate the ratio of income and price elasticity and a lower (for yes responses) or upper bound (for no responses) on individuals’ implicit WTP. If the survey asks the same questions about several alternative expenditure levels and quantities of the service, then it is possible to estimate demand curves (Lankford, 1985).

### 3.3 Budget Allocation Surveys with Fixed Total Budget and Actual Prices

Budget allocation surveys (or budget games) transfer individual decisions on a competitive market to a simulated market for public services. Rather than asking respondents about their demand for a single public service, the respondents are given a hypothetical budget which they can allocate to different public services such as public education or improvements of local environmental quality. The (actual) marginal prices of the individual services or projects are provided explicitly. The respondents may revise their allocations until no further changes are desired. As in a competitive market, the individuals reach their utility maximum when the price ratio of any two services equals the relative valuation of those services. Pendse and Wyckoff (1976) used a budget game to examine the marginal valuation of an additional river dam and also to derive which aspects such as flood control, protection of the natural landscape or water sport opportunities were particularly highly valued by the local residents. Denzau et al. (1977) applied this approach to six public services in the city of Tucson, Arizona. Furthermore, the authors suggested that the approach could be extended by allowing participants to change also the total public budget (see following section).

### 3.4 Budget Allocation Surveys with Variable Total Budget and Experimental Variation of Prices

To obtain not only relative valuations based on the existing total budget, the budget allocation survey approach can be extended by changes in the total budget (to examine reactions to changes in income), by allowing respondents to adjust the total budget or by experimentally varying the relative prices of the services in order to estimate demand functions or WTP for individual services. Strauss and Hughes (1976) examined the demand for twelve public services in North Carolina. Hardie and Strand (1979) applied the approach – with different price vectors – to estimate demands for different types of park area in a national park system. Hockley and Harbour (1983) applied the approach to public services in England and Wales. Schokkaert (1987) used a large survey to examine the demand for 24 public projects in a Belgian town.4
3.5 Contingent Valuation

In contingent valuation surveys, individuals are asked in self-contained surveys whether they prefer a proposed change in public service to (usually) the status quo. Various question formats have been proposed. However, to achieve more desirable incentive properties and questions that are easier to handle, a preferred format is the ‘referendum format’ in which the alternative policy is presented together with a price tag (or ‘bid’). The intended statistical analysis requires that the specified price is a randomly assigned amount and not what the policy would actually cost the respondent if it was implemented. Price figures are assigned according to an experimental design for the purpose of statistical identification of WTP distributions. Typically, the figures are varied over a wide range of values such as from $5 to $1000. The survey approach thus allows estimating distributions of stated WTP among samples of the population.

The initial idea for the approach goes back at least to Bowen (1943) who writes: ‘It is conceivable, moreover, that the voters might be asked to indicate their preferences at each of several possible prices, so that the [average marginal rate of substitution] could be ascertained along several points and the intersection of the [marginal rate of substitution and costs curves] could be located immediately’ (p. 40). One of the first empirical applications of the approach that was accepted in an economics journal was Randall et al. (1974), in the first volume of the Journal of Environmental Economics and Management. Applications increased greatly during the 1990ies due to the Exxon Valdez oil spill and subsequent research related to damage assessment and litigation.

3.6 Contingent Valuation with Experimental Variation of Percentage or Per-Unit Charges

In this approach, as in standard contingent valuation, the choices presented to the respondents involve randomly assigned prices drawn from an experimental design. However, the prices are specified as changes in tax rates or per-unit charges rather than as total individual amounts. For instance, if the policy is to be financed through an income tax, the price is specified as a percentage increase in that tax. A relatively narrow range of price (or ‘bid’) variation is then sufficient to cover a range of realistic individual payments (cf. Section 2.2).

As in standard contingent valuation, WTP distributions can be derived. However, the WTP units are tax percentage points or (tax) prices per unit of service. WTP distributions in monetary units must be derived in an additional step of the analysis. The approach has been applied in a number of recent studies (Schläpfer and Schmitt, 2007; Schläpfer et al., 2008; Rheinberger, 2011).

Key characteristics of the 11 survey approaches and empirical examples are summarized in Table 2.

4. Challenges Addressed by the Survey Approaches

4.1 Overview of Challenges

Measuring preferences for public goods poses a series of challenges, and specific strengths of alternative survey approaches may depend on whether and how they address the challenges (Table 3). Important challenges are: (i) cognitive limitations of the respondents, (ii) issues with incentives for answering strategically, (iii) (legitimate) effects of costs distribution on WTP for public services and (iv) an interest in measuring WTP distributions. In the following these challenges are briefly discussed.

4.2 Cognitive Limitations or Bounded Rationality

Bounded rationality (Kahneman, 2003) or a lack of ‘articulated values’ (Fischhoff, 1991) is a key challenge in stated preference research. Some of the existing approaches directly address this challenge. The various proposed solutions to the challenge are not mutually exclusive.
Table 2. Stated Preference Approaches for Public Services and the Preference Information they Produce.

| Survey approach                                      | Price information | Construct measured     | Question format                                                                 | Preference information produced                           | Examples: self-contained surveys | Examples: surveys with information on political support |
|------------------------------------------------------|-------------------|------------------------|-------------------------------------------------------------------------------|------------------------------------------------------------|---------------------------------|-------------------------------------------------------|
| Micro-based demand survey                           | Actual price      | Direction of quantity demanded | Preferred direction of quantity change given true price                       | Demand functions, Demand elasticity                        | Bergstrom et al. (1982)           | No example available                                   |
| Referendum survey                                    | Actual price      | Quantity demanded      | Choice between two alternatives given actual price (single or multiple projects) | Frequency of WTP values above/below costs; distribution of marginal WTP<sup>a</sup> | No example available            | Vossler and Kerkvliet (2003)                           |
| Budget allocation survey with actual prices          | Actual price      | Quantity demanded      | Preferred quantity given actual price                                         | Marginal rate of substitution between public goods         | Denzau et al. (1977)              | No example available                                   |
| Budget allocation survey with experimental variation of percentage or per-unit charges | Randomly assigned but realistic price | Quantity demanded      | Preferred quantity given randomly assigned per-unit prices (but actual cost distribution) | Marginal willingness-to-pay                                | Strauss and Hughes (1976)         | No example available                                   |
| CV survey with experimental variation of percentage or per-unit charges | Randomly assigned but realistic price | WTP                    | Choice between two alternatives given randomly assigned per-unit prices (but actual cost distribution) | Distribution of marginal WTP (in monetary units or in units of the relevant tax base) | Rheinberger (2011)                  | Schläpfer et al. (2008)                             |
| Standard CV survey (with experimental variation of total individual costs) | Randomly assigned and unrealistic price | WTP                    | Choice between two alternatives given randomly assigned individual costs       | Distribution of WTP (in monetary units)                    | McFadden (1994)                   | Not feasible; see Table 1                              |

<sup>a</sup>The distribution can be obtained by asking about multiple projects with different size if marginal WTP is constant and marginal cost increasing.
### Table 3. Challenges Addressed by Different Survey Approaches.

| Survey Type                                      | Self-contained survey                                                                 | Survey with information on political support                                      |
|--------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| Micro-based demand survey                        | Scenario credibility and incentives; cognitive limitations; relevance of cost distribution | Scenario credibility and incentives; cognitive limitations; relevance of cost distribution |
| Referendum survey                                 | Scenario credibility and incentives; relevance of cost distribution                    | Scenario credibility and incentives; cognitive limitations; relevance of cost distribution |
| Budget allocation survey with actual prices       | Scenario credibility and incentives; relevance of cost distribution                    | Incentive issues; cognitive limitations; relevance of cost distribution               |
| Budget allocation survey with experimental variation of percentage or per-unit charges | Relevance of cost distribution; interest in WTP distributions                         | Cognitive limitations; relevance of cost distribution; interest in WTP distributions |
| CV with experimental variation of percentage or per-unit charges | Scenario credibility and incentives; relevance of cost distribution; interest in WTP distributions | Scenario credibility and incentives; cognitive limitations; relevance of cost distribution; interest in WTP distributions |
| Standard CV (with experimental variation of total individual costs) | Interest in WTP distributions                                                       | Not feasible; see Table 1                                                            |

One solution is to ask only qualitative questions about the preferred direction of marginal changes. Bergstrom et al. (1982, p. 1186) write: ‘Asking a simple qualitative question about the direction of the respondent’s preferred amount of public expenditure from the status quo, rather than asking him to specify more exactly how much he would like, reduces the burden on the respondent’s imagination’.

This is the approach used in (self-contained) micro-based demand surveys or in referendum decisions about small changes in provision. Especially if questions involve only major categories of (existing) public services, it seems reasonable to assume that many respondents will know from past political debates whether they prefer rather more or rather less of the services.

A second solution is to reduce the cognitive burden by asking questions involving actual prices rather than counterfactual prices. This approach is followed in qualitative demand surveys, in referendum surveys (including actual referenda), and sometimes in budget allocation surveys. In politics, actual prices may reduce the cognitive burden since for a given level of efficiency in provision they may convey information about the quantity or quality of a service that may be easily understood and possibly compared with the costs of earlier projects.

Third, questions about expenditures for (or quantities of) single services may be presented in the context of other budget categories and total budgets. In this way, respondents are provided with potentially helpful reference information. For instance, WTP for maintaining bird habitat on agricultural lands may be easier to estimate if the budgets for other agri-environmental objectives and the total agri-environmental budget is
known. This solution is followed in budget allocation surveys. Potential problems with bounded rationality such as scope insensitivity or anchoring effects (e.g. McFadden, 1994), may thus be decreased.

Finally, questions may be accompanied by information on the distribution of political support of the proposed services by better informed individuals or organizations such as political parties or interest groups. Such information may be accessible through the media and public debates or it can be specially provided by the researcher (see Section 2.3).

4.3 Scenario Credibility and Incentive Compatibility

A second issue in stated preference research on public services is related to incentives for answering strategically. This issue is not confined to open-ended questions about WTP for public goods but applies also to dichotomous questions if there are more than two alternatives or if the alternatives are not the true and only alternatives (Gibbard, 1973; Satterthwaite, 1975). The latter situation occurs in conventional contingent valuation due to the randomly assigned prices or ‘bids’ in dichotomous questions. Since these prices are not the true prices if a proposed policy is implemented (Champ et al., 2002) the respondents who do not find the prices credible may understand that they have the opportunity to answer strategically. Specifically, if a person’s WTP for a public service is higher (lower) than his or her actual expected tax consequence, this person should always accept (reject) the policy, regardless of the price in the (dichotomous) CV question.

The simplest way to solve the problem is by only asking questions in a form that reflects the Gibbard–Satterthwaite result. This is the case in actual referendum questions or in referendum surveys (e.g. Vossler et al., 2003), provided the respondents believe that the status quo and the alternative are the true and only options. The problem is similarly solved in qualitative demand surveys where respondents only state whether they prefer more, less or the same amount of the public services.

If randomly assigned prices are needed to identify WTP distributions, then a solution is to keep the experimental variation of the price within a narrow range around the actual (or expected) price. This mitigates the problem of prices that may not be credible and the resulting possibility that the respondents answer strategically or answer different questions than those intended by the researcher (Flores and Strong, 2007). An effective way to keep the randomly assigned price within a narrow (but relevant) range around the actual price is then to formulate the price as a percentage or per-unit charge (see Section 4.5. below).

4.4 Effects of Cost Distribution on WTP

Another issue that troubles stated preference research is the potential relevance of the distribution of policy costs (Cai et al., 2010). If preference estimates are sensitive to the distribution of costs, it will be important how this distribution is specified.

One obvious solution to this challenge is specifying in the survey question the actual (and hence relevant) distribution of costs. In this way the respondents whose choices are sensitive to the financing arrangement can base their choices on the actual and relevant distribution of costs. This is followed in approaches that specify actual costs or that specify costs as realistic percentage or per-unit charges (cf. Table 1).

4.5 Interest in Estimating WTP Distributions

From an applied welfare perspective it is useful if a stated preference approach provides entire WTP distributions for public services rather than only upper or lower bounds as in the economic analysis of individual referendum choices. To achieve this, the prices in referendum questions about WTP have to be experimentally assigned (Hanemann, 1984). Otherwise it is not possible to statistically isolate WTP distributions (see Section 2.2).
Table 4. Overview of Critical Assumptions of the Survey Approaches.

| Survey Type                                | Assumptions                                                                 |
|--------------------------------------------|-----------------------------------------------------------------------------|
| Self-contained survey                      | Motivation; knowledge about direction of preferences                         |
| Survey with information on political support| Motivation; efficient use of advice                                           |
| Micro-based demand survey                  | Motivation; individual rationality                                           |
| Referendum survey                          | Motivation; efficient use of advice                                           |
| Budget allocation survey with actual prices| Motivation; individual rationality; no strategic answers                      |
|                                            | Motivation; efficient use of advice                                           |
| Budget allocation survey with experimental variation of percentage or per-unit charges| Motivation; individual rationality; no strategic answers                      |
|                                            | Motivation; no strategic answers; efficient use of advice                     |
| CV survey with experimental variation of percentage or per-unit charges| Motivation; individual rationality                                           |
|                                            | Motivation; efficient use of advice                                           |
| Standard CV (with experimental variation of total individual costs) | Motivation; individual rationality; no strategic answers; no sensitivity to cost distribution |

The solution in contingent valuation is to randomly assign prices at the level of the individual respondent. Another solution, used in budget allocation surveys and in a variant of contingent valuation, is to use experimental variation of the relevant percentage or per-unit charges. This can be percentage changes in taxes or per-unit changes in utility charges. The advantage of this approach is that distributions of WTP (within a relevant range of WTP) can be assessed without randomly assigned, counterfactual cost figures (cf. Section 3.6).

5. Required Assumptions of the Approaches

5.1 Overview of Assumptions

As a consequence of the specific sets of challenges addressed by the different stated preference approaches (Table 3) there are different sets of critical assumptions on which the preference elicitation relies (Table 4). Important relevant assumptions concern (1) the respondents’ motivation to make a serious effort to answer the questions, (2) the capacity to answer the questions in line with personal interests and values in spite of cognitive limitations, (3) no strategic answering (if opportunities exist) and (4) no sensitivity to cost distribution (if cost distribution in the survey is unspecified or counterfactual). In the following, these assumptions are briefly discussed.

5.2 Respondent Motivation

A sufficient motivation to make a serious effort at answering the questions is an obvious requirement in any stated preference approach. While motivation in actual voting decisions is determined by various
personal and contextual factors (e.g. Kriesi, 2007), motivation in sample surveys may also depend heavily on how the respondents perceive the likely importance and impact of the survey on public policy (Carson and Groves, 2007). If the respondents perceive the survey to be consequential, they may also be motivated by a desire to influence the outcome of the survey in their favour.

5.3 Capacity to Answer in Line with Personal Interests and Values

The assumptions regarding a sufficient capacity to answer in line with personal interests and values take various forms – from rationality in a narrow economic sense (as assumed in conventional CV) to ‘knowledge about the preferred direction of marginal changes of the public service’ (Bergstrom et al., 1982) and to ‘efficient use of advice’, as required in approaches in which information on the political support of proposed policies is available (Lupia, 1994). While media exposure and experience from political debates may support knowledge about the preferred direction (or marginal changes) of public expenditure changes, these sources of information obviously cannot inform individual preferences regarding discrete changes in public services that have never been publicly discussed. If survey respondents can access political advice from better informed individuals or interest groups, the assumption of individual rationality is replaced by an assumption of ‘efficient use of advice’. Individuals are assumed to understand whose political advice they can trust in identifying their preferred choice (Druckman, 2001; Lupia and Matsusaka, 2004, Bechtel et al., 2015). Little is currently known, however, about the conditions under which each of the alternative assumptions is most appropriate.

5.4 No Strategic Answering

Those survey approaches which make use of randomly assigned prices – rather than true or at least credible prices – further require the assumption of ‘no strategic answers’. More precisely, following the explanations in Section 4, the assumption is required if questions about WTP or quantities demanded are open-ended or if – in dichotomous questions – the alternatives (including their prices) are not the true and only alternatives (the Gibbard–Satterthwaite result). However, if there are only two alternatives, and these alternatives are realistic enough such that respondents cannot know they are not the true and only alternatives, then the respondents do not perceive any incentive to answer strategically. This is the situation of (dichotomous-choice) budget allocations and CV surveys with experimental variation of percentage or per-unit charges.

5.5 No Sensitivity to Cost Distribution

As mentioned in the Section 4, the sensitivity of responses to cost distribution may be addressed by presenting unambiguous information about actual cost distribution in the survey. Where this is not done – due to an inherent conflict between realistic payment vehicles and randomly assigned individual costs (‘bids’) in the survey questions – the survey analysis requires an assumption that cost distribution does not matter.

6. Evaluation of Approaches in Light of Empirical Evidence and Theoretical Considerations

6.1 Empirical Evidence

Much research on the validity of stated preference surveys has focused on standard contingent valuation applied to private goods, quasi-public goods, donations to public goods, or group donation mechanisms (see Journal of Economic Surveys (2017) Vol. 31, No. 1, pp. 258–280 © 2015 John Wiley & Sons Ltd
Kling et al., 2012 for a summary of validity concepts and an overview of the evidence for contingent valuation). Unfortunately, these studies reveal limited insights about the validity of surveys about those types of unfamiliar public goods in which applied research is typically interested. The evidence on the validity of stated preference surveys for those goods is fairly limited even in the case of standard contingent valuation. The most relevant empirical approaches are assessments of ‘construct validity’ examining whether the estimates are appropriately (in)sensitive to changes in the questions and assessments of ‘criterion validity’ examining whether the estimates are consistent with those observed in actual voting decisions.

Powerful experimental tests of construct validity have been the exception rather than the rule in applied work, which may be surprising given economists’ traditional skepticism towards stated preference methods. For most applied work we have no means to know how large any measurement error may be. Experimental tests of survey validity are available for standard contingent valuation applications. However, the test most frequently applied – the ‘scope’ test which examines the sensitivity of the estimates to the amount or scope of the public good – provides only limited evidence, since it is not clear how sensitive the estimates should be (Hanemann, 1994, p. 35; McFadden, 1994, p. 702; Desvousges et al., 2012). Potentially powerful anchor tests and other framing manipulations have demonstrated that CV estimates for public goods (e.g. Green et al., 1998) and even for quasi-public goods (Hausman, 2012) can be highly sensitive to theoretically irrelevant changes in response scales or survey wording. However, apart from the study mentioned in Hausman (2012), such tests have not been implemented in state-of-the-art applied research.

The criterion validity of surveys about public goods – whether stated values match real payments – has been assessed through comparisons with actual voting behaviour. Such comparisons are available only for referendum surveys where the evidence suggests a high degree of validity (e.g. Vossler and Kerkvliet, 2003; Vossler et al., 2003; Johnston, 2006) and for standard contingent valuation where the (limited) evidence is less encouraging (Shabman and Stephenson, 1996; Schläpfer et al., 2004; Schläpfer and Hanley, 2006). Evidence for contingent valuation with credible prices and political advice suggests that those valuations may be more consistent with individual interests and values than contingent valuation with credible prices without advice (Schläpfer et al., 2008; Schläpfer, 2011).

Much remains to be done in assessing the validity of alternative survey approaches and in interpreting the related evidence (Kling et al., 2012; Haab et al., 2013). Due to the lack of further empirical evidence on the performance of the other approaches, the discussion of relative advantages and disadvantages of those approaches must heavily rely on theoretical considerations.

6.2 Micro-Based Demand Surveys

An important advantage of micro-based demand surveys which use only qualitative questions is that issues with cognitive limitations may not loom as large as in quantitative questions about the demand or WTP for public goods. At least for major classes of public services, most individuals may know from personal experience whether they would benefit from marginal increases or decreases in provision, given the existing tax structure or other relevant public finance system. Furthermore, issues with scenario credibility are unlikely to arise, and there are no opportunities for answering strategically. A disadvantage is that the approach only allows researchers to compute demand functions or demand elasticity near current levels of service provision. In principle, this restriction could be relaxed and respondents could be asked quantitative questions about their demand for alternative levels of service. However, such a move would also bring back issues with cognitive limitation and incentive compatibility that are resolved by the original approach.

6.3 Referendum Surveys

A key strength of referendum surveys is that they allow researchers to present true and unambiguous policy alternatives and thus to provide as much detail as seems manageable and useful for competent
choices. Further potential benefits of presenting the true policy alternatives are the absence of incentives for answering strategically and the possibility to address issues with bounded rationality and unfamiliar services by providing political expert advice.

The disadvantage is that it is usually not possible to estimate entire WTP distributions (but see Section 3.2). It is possible to estimate the plurality of the respondents with a WTP larger than their expected costs. The sum of the expected costs of the approving respondents is then a conservative lower bound estimate of aggregate WTP (Schläpfer and Hanley, 2006).

6.4 Budget Allocation Surveys

Budget allocation surveys are an intuitively meaningful approach to determine the relative valuation of public services. An important advantage is that they may reduce biases that occur when single public goods are valued in isolation (e.g. McFadden, 1994; see Section 4.2). For local public services and near current levels of provision, respondents can be expected to know from experience which way they would like to shift the expenditures for various services. This is especially true if the total budget is fixed, and hence, the questions involve only changes in relative valuations among service categories. In that case, the situation is similar to qualitative surveys about quantity demanded (cf. Section 6.2).

When the prices of public services are manipulated, or when the total budget can be adjusted by the respondents, the choice tasks become increasingly complex, and choices must be expected to become more uncertain. Nevertheless, as long as the alternative levels of provision involve realistic cost distributions (e.g. tax rates), the cognitive demands on the respondents remain lighter than in the standard contingent valuation approach where prices do not carry additional information about service quantity or quality (cf. Section 4.2).

6.5 Contingent Valuation

The key strength of contingent valuation with randomly assigned costs is its potential to explore entire WTP distributions including individual WTP values at the tails of the distribution. The approach would be ideally suited for all applied purposes if concerns about scenario credibility, distribution of costs, strategic answering and cognitive limitations did not arise. Interestingly, the characteristic of counterfactual prices or ‘bids’ in contingent valuation questions about public goods have received little attention in the literature (but see Flores and Strong, 2007). The issue with counterfactual prices is the incompatibility of two key requirements in contingent valuation: On the one hand, informed valuation requires that the survey instrument provide accurate information about how the policy will be financed (e.g. Arrow et al., 1993; Kling et al., 2012). As a consequence the respondents are able to form expectations about their individual tax implications. On the other hand, the intended analysis requires that the bid values are randomly assigned. Due to their random assignment, these amounts will in general conflict the tax expectations formed on the basis of proposed financing arrangements (cf. Section 2.2). This issue constitutes a fundamental dilemma of the standard contingent valuation approach which may undermine what is sometimes called the ‘face validity’ of the survey instrument (cf. Carson, 2012, p. 39).

Due to the random assignment of prices is also impossible to provide political advice to the individual respondents. Political parties and interest groups will find it impossible to define their position on questions involving counterfactual (and conflicting) information on financing arrangements and individually assigned costs.

In sum, the reliance on randomly assigned, counterfactual prices conflicts fundamentally with a precise description of the alternative scenarios. And since the precise description of the alternatives is the basis for measuring economic values, the potential for improving the validity of standard contingent valuation seems to be fundamentally limited. It is noteworthy in this context that some of the strongest evidence
of CV validity discussed in Kling et al. (2012, p. 16) – the comparisons with actual voting outcomes – involves SP questions using *actual* rather than randomly assigned prices and settings where information about the distribution of political support is available. Following the present classification, that survey approach is a referendum survey with information on political support (c.f. Section 6.3).

6.6 Contingent Valuation with Experimental Variation of Percentage or Per-Unit Charges

Contingent valuation with randomly assigned percentage or per-unit (changes) in taxes or other payments combines the advantages of both randomly assigned prices and realistic cost distributions. The approach allows the researcher to provide sufficiently precise information about the financing mechanism and to formulate the individual price as a realistic and hence credible percentage change in the relevant tax. The respondents’ absolute tax amounts are calculated based on the individuals’ tax factors. In a survey instrument, a table can be provided in which respondents with different tax factors may look up their personal tax contribution if the policy is approved.

CV with realistic price distribution largely eliminates credibility problems since the individually assigned values remain within a credible range (cf. Section 2.2). The surveys may be supplemented with expert advice or other information on the political support of the alternatives which may help address issues with cognitive limitations (see following section). Compared with traditional CV, one drawback is that the analyst obtains WTP in tax percentage units. These units must be converted to dollar units, taking individual tax factors into account.

6.7 Alternatives in the Provision of Policy Information

Self-contained surveys of stated preferences for discrete changes in public goods require strong assumptions about the respondents’ cognitive abilities and knowledge of their preferences (see Table 2). The assumption of ‘known preferences’ and unrealistic cognitive abilities can be dropped when advice from trusted experts is made available in a survey. Such information can be available from a public discussion – if the specific issue of the survey is already a matter of public debate – or it can be gathered specifically for the purposes of the survey. However, information from trusted experts can only be provided to the respondents in surveys that use realistic price distributions. In surveys using randomly assigned bids (the case of standard contingent valuation), it is not possible for experts to provide useful advice, since individual payment levels conflict with the general specification of the payment mechanism (see Section 6.5). Advice from trusted experts has the potential to make preference elicitation more reliable in situations where respondents do not have prior choice experience.

A drawback of such approaches is that information about the distribution of political support may not only help respondents identify their preferred choices but could also influence responses in directions that do not agree with the respondents’ underlying interests and values. To what extent this happens is relevant from a welfare economic perspective and should be evaluated empirically. Providing information about the distribution of political support can be viewed as an example of libertarian paternalism (Sunstein and Thaler, 2003) where, in this case, the issue position of a respondent’s favoured party or interest group has the character of a respondent-specific default option. As in other applications of libertarian paternalism, more empirical work is needed to further evaluate potential merits and welfare implications of the approach.

Again (cf. Section 6.5), it is worth noting that some of the best evidence for CV validity from comparisons with actual voting outcomes cited in recent assessments (Murphy et al., 2005; Kling et al., 2012) involves SP questions that were asked in a context where information about the distribution of political support was available (e.g. Vossler and Kerkvliet, 2003). The approach in those surveys is not a standard contingent valuation survey – and this is acknowledged by the authors (p. 647) – but the referendum survey approach in the upper-right cell of Table 1.
Table 5. Applied Potential of the Survey Approaches.

| Method                                                                                       | Self-contained survey                                                                 | Survey with information on political support                                                                 |
|--------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| Micro-based demand survey (in sample of districts)                                          | –Measure demand elasticities                                                           | –Measure demand elasticities                                                                               |
| Referendum survey                                                                           | –Measure the plurality supporting a (realistic) proposed policy if the public service is sufficiently familiar | –Measure the plurality supporting a (realistic) proposed policy                                            |
|                                                                             | –Estimate bounds on mean WTP in cases where limited information is sufficient for cost-benefit analysis | –Measure ‘political WTP’ (approved budget given the existing tax schedule or other financing arrangement) |
| Budget allocation survey with fixed total budget and actual prices                          | –Measure relative valuations (marginal rates of substitution) of groups of public services if the public good is sufficiently familiar | –Measure relative valuations (marginal rates of substitution) for groups of public services               |
| Budget allocation survey with variable total budget and experimental variation of percentage or per-unit charges | –Measure valuations for groups of public services if the public good is sufficiently familiar | –Measure valuations for groups of public services                                                          |
| CV with experimental variation of percentage or per-unit charges                            | –Measure mean WTP for a specific policy if the public good is sufficiently familiar       | –Measure mean WTP for a specific policy                                                                   |
|                                                                             | –Estimate bounds on mean WTP in cases where limited information is sufficient for cost-benefit analysis |                                                                                |
| Standard CV (with experimental variation of total individual costs)                         | –Examine entire WTP distributions for public goods                                      | Not feasible; see Table 1                                                                                 |
|                                                                             | –Estimate mean WTP where a relatively large degree of uncertainty is tolerable          |                                                                                |

7. Applied Potential of the Alternative Approaches

7.1 Overview of the Approaches’ Applied Potential

Table 5 provides an overview of the applied potential of each approach based on their required assumptions and the specific challenges they address. The approaches’ relevant fields of application are not mutually exclusive. There are several applied purposes where the results from one or several alternative approaches could be used in addition to the currently dominant point estimates of maximum WTP as obtained in self-contained contingent valuation with randomly assigned prices. In some decision making contexts other preference parameters such as the median preferred budget may be more directly useful to policy
makers than the mean or median WTP for a specific policy. In the following, the approaches available for various applied purposes are briefly discussed.

7.2 Demonstrate that a Proposed Public Good is Valued

If the goal of the study is to demonstrate that a public good is valued by many people, relatively uncertain values which rely on restrictive assumptions, such as familiarity with the public good or inattention to opportunities for answering strategically, may be sufficient. It seems fair to say that for some applications of contingent valuation by lobby groups or government agencies the main objective is not to estimate values as accurately as possible but to demonstrate a substantial WTP for implementing a given project. An (average quality) standard contingent valuation survey without powerful validity tests may be appropriate for this purpose in spite of the considerable bias reported in the literature (Kling et al., p 15). This is not to say, of course, that average standard contingent valuation surveys should remain on this level of accuracy in the future.

7.3 Explore WTP Distributions where a Relatively High Degree of Uncertainty is Tolerable

The main applied potential of the present state-of-the-art in contingent valuation may be to provide estimates of maximum WTP in cases where entire WTP distributions are required but a relatively large degree of uncertainty is tolerable. This at least seems to be a conclusion from the assessments in Arrow et al. (1993), List and Gallet (2001), Murphy et al. (2005), or Kling et al. (2012). Based on the preceding theoretical evaluation, CV with randomly assigned but realistic prices or CV with randomly assigned but realistic prices plus information on political support of the proposed policies (see Table 1) may be appropriate for the same applied purpose. In addition, where rough upper or lower bound estimates may be sufficient (see e.g. Kling et al., 2012, p. 17), referendum surveys with actual price information as previously used in CV validation (e.g. Vossler and Kerkvliet, 2003) may play a similar role (cf. Section 7.6).

7.4 Estimate WTP for a Realistic Policy Scenario as Accurately as Possible

Measuring maximum WTP for a realistic policy as accurately as possible requires the use of randomly assigned prices. However, the random assignment of prices should not preclude the use of realistic information on cost distribution (tax vehicles and individual payments). Furthermore, access to information about the support of the proposed policies may be appropriate in some cases to help respondents make decisions about novel and unfamiliar public issues in line with their interest and values. Careful elicitation of politically informed preferences in contingent valuation questions with experimental variation of percentage or per-unit charges may turn out to be an appropriate survey approach. While the evidence on this approach is still limited, the predictive success of pre-election surveys suggests that surveys with realistic price ranges should be appropriate for a fairly accurate measurement of values.

In special cases where marginal WTP is known to be constant and (known) costs increasing, preference questions with true prices may be used to infer WTP for the public good. More specifically, marginal WTP (and hence total WTP for a quantity change) could be inferred from responses to questions about preferred quantities given the (true) cost curve (see Section 3.2).

7.5 Estimate the Median Preferred Budget and Quantity

If the goal is to estimate the preferred quantities of various services given the existing means of financing, for example in the agri-environment domain, then budget allocation surveys with true prices or
micro-based demand surveys with quantitative questions about preferred levels may be more appropriate in some cases than WTP estimates for specific quantities. In the case of unfamiliar goods, external information on political support of the proposed policies may be useful to avoid the pitfalls of an unrealistic assumption that individuals are able to make choices about complex public services in line with their interests and values. Potential advantages of the additional information, however, would have to be weighed against concerns that the additional information may not always influence the respondents’ choices in the direction of their underlying preferences.

7.6 Estimate the Proportion of the Population that Wants a Policy to Go Forward

If the goal is to estimate what proportion of the population feels that a proposed service is worth what it will cost the individual taxpayer or consumer, given the relevant means of financing the associated expenditure, then simple referendum survey questions with true prices may be useful. As mentioned in Kling et al. (2012, p. 17), simple upper or lower bound estimates of passive use value can sometimes be sufficient for a cost-benefit test. Again, additional information on the distribution of political support as available in actual referenda may be useful to avoid unrealistic assumptions on the respondents’ cognitive abilities.

8. Discussion and Conclusions

Stated preferences research until the 1990s has produced a variety of alternative approaches to measure preference parameters relevant to decisions about the provision of public services, and each one of them has its specific strengths. The recent literature does not reflect this diversity. One single approach, contingent valuation (including choice experiments), seems to have become the method of choice whenever revealed preference information is unavailable. The dominant role of contingent valuation does not appear to be the result of a systematic evaluation of the alternatives for different objectives. The advantages and disadvantages of alternative approaches have hardly been examined. In particular, although public voting decisions are widely perceived as a blueprint for survey design and a standard in validation (see e.g. Kling et al., 2012), little attention has been paid to the dimensions in which contingent valuation and other SP surveys differ from voting decisions. Two of these dimensions – the nature of the price information and the presence/absence of information about the distribution of political support of the policy – were used here to classify and discuss a range of approaches that use sample surveys to measure stated preferences for public goods. This classification yielded eleven available alternatives in stated preference research which follow different rationales and use different critical assumptions.

The classification proved useful in organizing the approaches, since it turned out to be closely linked with different preference parameters produced and different assumptions required. The use of actual versus randomly assigned prices determines whether or not WTP distributions can be estimated and also implies different assumptions regarding respondents’ perception of ‘prices’ in surveys (Champ et al., 2002; Flores and Strong, 2007). The use of self-contained versus nonself-contained surveys implies different assumptions on individual rationality – in a narrow economic sense versus in the sense of an ‘ecological rationality’ based on a competent use of heuristics in a political context (Smith, 2003; Lupia and Matsusaka, 2004). The classification thus helped to identify unique combinations of critical assumptions, which in turn implied specific strengths and limitations for different applied purposes.

The findings suggest that different approaches may be preferable for different applied purposes. Simple qualitative preference questions following Bergstrom et al. (1982) are well suited to estimate demand parameters near current levels of provision and where cross-sectional variation in current provision can be exploited. Incentive-compatible referendum questions involving true cost estimates may be ideally suited if the main research objective is to examine public support for the funding of a proposed service.
when a credible and transparent specification of the alternatives is important. Budget allocation surveys seem particularly useful for research interested in relative valuations across a range of public services within an existing public budget context. WTP questions involving realistic prices drawn from relevant pricing schemes such as given tax schedules seem appropriate if credible scenarios and thus accurate and policy-relevant WTP estimates are sought. Survey approaches providing respondents with information about the distribution of political support of the proposed policies may sometimes be appropriate in the case of unfamiliar public services. Finally, the standard contingent valuation approach with its use of counterfactual prices which are randomly drawn from a wide range of bid values seems ideally suited if the objective is to explore entire WTP distributions.

As the existence of environmental values is now widely accepted, the interest of stated preference research is likely to shift from the demonstration of value towards the accurate measurement of value. Unfortunately, evidence from powerful validity tests is extremely rare. Even for CV, the evidence is fairly limited since the scope tests often used in state-of-the-art surveys provide only weak information (Hanemann, 1994; Desvousges et al., 2012; Kling et al., 2012). The arguably strongest validity tests – framing experiments such as anchoring experiments – have been used in some contingent valuation surveys (e.g. Green et al., 1998) but hardly ever in applied research on public goods. Validity tests using public voting decisions as a criterion are available mainly for referendum surveys using actual prices and informational environments similar to those in actual votes (e.g. Vossler and Kerkvliet, 2003).

Commentators of contingent valuation have suggested that contingent valuation may be 'the only game in town' when it comes to estimation of passive use value (e.g. Kling et al., 2012, p. 4), and the alternative to contingent valuation would be 'to place a zero value on goods that the public cares about' (Carson, 2012, p. 40). That logic suggests a search for further improvements of stated preference valuation within the standard contingent valuation paradigm (e.g. Carson, 2012, p. 40; Kling et al., 2012, p. 22). The evaluation in this paper suggests that a broader perspective may help identify useful additional directions. The findings strongly suggest that each approach has something to offer, and there is no single best approach. Each approach strikes a balance between the partly conflicting goals of measuring entire WTP distributions and asking realistic, credible, incentive compatible, manageable questions. The strengths offered by the examined alternatives therefore relate to the specific ways in which they resolve potential issues with scenario realism and credibility, incentive compatibility, and cognitive limitations (see Section 7).

Future progress in the search of appropriate stated preference methods for specific applied purposes may benefit from a more systematic assessment of validity. In survey research by psychologists and behavioural economists over the past decades, anchoring experiments and other framing experiments have been identified as a particularly powerful standard for that assessment (e.g. Jacowitz and Kahneman, 1995; Green et al., 1998; McFadden, 2001; Ariely et al., 2003). A wider use of standardized anchoring experiments in particular would not only facilitate the comparative evaluation of the existing survey approaches but also promote a fruitful and exciting competition for survey quality.

Acknowledgements
Detailed comments by two anonymous reviewers are gratefully acknowledged. They helped to substantially improve the manuscript.

Notes
1. In the following the term contingent valuation (CV) always includes the choice experiment (CE) approach.
2. An exception is the German-language monograph on preference elicitation by Pommerehne (1987) which, however, does not provide an empirical assessment.

3. If it may be reasonably assumed that the marginal WTP for a policy is constant while marginal costs are increasing in a relevant range – as for example in the WTP for saving additional human lives – then survey questions about multiple levels of the public good may provide information about the distribution of marginal WTP.

4. A further option to estimate absolute values is to insert an externally derived marginal WTP value for one of the services included in the budget game (Pommerehne, 1987, p. 188).

5. To facilitate the conversion of the percentage tax increase into dollar amounts, a table may be provided from which the respondents may look up the tax consequences, given the relevant tax structure and income bracket.

References

Ahlin, A. and Johansson, E. (2001) Individual demand for local public schooling: evidence from Swedish survey data. *International Tax and Public Finance* 8(4): 331–351.

Alevy, J.E., List, J.A. and Adamowicz, W.L. (2011) How can behavioral economics inform nonmarket valuation? An example from the preference reversal literature. *Land Economics* 87(3): 365–381.

Ariely, D., Loewenstein, G. and Prelec, D. (2003) "Coherent arbitrariness": stable demand curves without stable preferences. *Quarterly Journal of Economics* 118(1): 73–105.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R. *et al.* (1993) Report of the NOAA panel on contingent valuation. *Federal Register* 58(10): 4601–4644.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N. *et al.* (2002) Economic valuation with stated preference techniques. A *Manual*. Cheltenham, U.K.: Edward Elgar.

Bechtel, M., Hainmueller, J., Hangartner, D. and Helbling, M. (2015) Reality bites: the limits of framing effects in salient policy decisions. *Political Science Research and Methods* 3(3): 683–695.

Bergstrom, T.C., Roberts, J.A., Rubinfeld, D.L. and Shapiro, P. (1988) A test for efficiency in the supply of public education. *Journal of Public Economics* 35(3): 289–307.

Bergstrom, T.C., Rubinfeld, D.L. and Shapiro, P. (1982) Micro-based estimates of demand functions for local school expenditures. *Econometrica* 50(5): 1183–1205.

Bowen, H.R. (1943) The interpretation of voting in the allocation of economic resources. *Quarterly Journal of Economics* 58: 27–48.

Brookshire, D.S. and Crocker, T.D. (1981) The advantages of contingent valuation methods for benefit-cost-analysis. *Public Choice* 36(2): 235–252.

Cai, B., Cameron, T. and Geoffrey, G. (2010) Distributional preferences and the incidence of costs and benefits in climate change policy. *Environmental & Resource Economics* 46(4): 429–458.

Carlsson, F. (2012) Design of stated preference surveys: is there more to learn from behavioral economics? *Environmental & Resource Economics* 46(2): 167–177.

Carson, R.T. (2012) Contingent valuation: a practical alternative when prices aren’t available. *Journal of Economic Perspectives* 26(4): 27–42.

Carson, R.T. and Groves, T. (2007) Incentive and informational properties of preference questions. *Environmental & Resource Economics* 37(1): 181–210.

Carson, R.T. and Louviere, J.J. (2011) A common nomenclature for stated preference elicitation approaches. *Environmental & Resource Economics* 49(4): 539–559.

Champ, P.A., Flores, N.E., Brown, T.C. and Chivers, J. (2002) Contingent valuation and incentives. *Land Economics* 78(4): 591–604.

Cummings, R.G., Harrison, G.W. and Rutstrom, E.E. (1995) Homegrown values and hypothetical surveys – is the dichotomous choice approach incentive-compatible. *American Economic Review* 85(1): 260–266.

Denzau, A.T., Weisz, R.N. and Walcott, J.C. (1977) A quasi-experimental method for studying public sector demands. *Southern Economic Journal* 44(2): 306–312.

Desvousges, W.H., Mathews, K. and Train, K. (2012) Adequate responsiveness to scope in contingent valuation. *Ecological Economics* 84: 121–128.
Diamond, P.A. and Hausman, J.A. (1994) Contingent valuation – is some number better than no number. *Journal of Economic Perspectives* 8(4): 45–64.

Druckman, J.N. (2001) Using credible advice to overcome framing effects. *Journal for Law, Economics & Organization* 17: 62–82.

Fischel, W.A. (1979) Determinants of voting on environmental-quality – study of a new-hampshire pulp-mill referendum. *Journal of Environmental Economics and Management* 6(2): 107–118.

Fischhoff, B. (1991) Value elicitation – is there anything in there? *American Psychologist* 46(8): 835–847.

Flores, N.E. and Strong, A. (2007) Cost credibility and the stated preference analysis of public goods. *Resource and Energy Economics* 29: 195–205.

Getzner, M. (2012) The regional context of infrastructure policy and environmental valuation: the importance of stakeholders’ opinions. *Journal of Environmental Economics and Policy* 1(3): 255–275.

Gibbard, A. (1973) Manipulation of voting schemes – general result. *Econometrica* 41(4): 587–601.

Gramlich, E.M. and Rubinfield, D.L. (1982) Micro estimates of public spending demand-functions and tests of the tiebout and median-voter hypotheses. *Journal of Political Economy* 90(3): 536–560.

Green, D., Jacowitz, K.E., Kahneman, D. and McFadden, D. (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics* 20(2): 85–116.

Haab, T.C., Interis, M.G., Petrolia, D.R. and Whitehead, J.C. (2013) From hopeless to curious? Thoughts on Hausman’s “dubious to hopeless” critique of contingent valuation. *Applied Economic Perspectives and Policy* 35(3): 593–612.

Hanemann, W.M. (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8(4): 19–43.

Hanemann, W.M. (1984) Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics* 66(3): 332–341.

Hardie, I.W. and Strand, I. (1979) Measurement of economic benefits for potential public goods. *American Journal of Agricultural Economics* 61: 313–317.

Hausman, J. (2012) Contingent valuation: from dubious to hopeless. *Journal of Economic Perspectives* 26(4): 43–56.

Hockley, G.C. and Harbour, G. (1983) Revealed preferences between public expenditures and taxation cuts: public sector choice. *Journal of Public Economics* 22: 387–399.

Jacowitz, K.E. and Kahneman, D. (1995) Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin* 21(11): 1161–1166.

Johnston, R.J. (2006) Is hypothetical bias universal? Validating contingent valuation responses using a binding public referendum. *Journal of Environmental Economics and Management* 52: 469–481.

Kahneman, D. (2003) Maps of bounded rationality. *American Economic Review* 93(5): 1449–1475.

Kahneman, D., Ritov, I. and Schkade, D.A. (1999) Economic preferences or attitude expressions?: an analysis of dollar responses to public issues. *Journal of Risk and Uncertainty* 19(1–3): 203–235.

Kling, L.K., Phaneuf, D.J. and Zhao, J. (2012) From exxon to BP: has some number become better than no number? *Journal of Economic Perspectives* 26(4): 3–26.

Kriesi, H. (2007) The participation in Swiss direct-democratic votes. In C.H. de Vreese (ed.), *The Dynamics of Referendum Campaigns* (pp. 117–141). Basingstoke: Palgrave Macmillan.

Lankford, R.H. (1985) Preferences of citizens for public expenditures on elementary and secondary education. *Journal of Econometrics* 27: 1–20.

Levinson, A. (2012) Valuing public goods using happiness data: the case of air quality. *Journal of Public Economics* 96: 869–880.

List, J.A. and Gallet, C.A. (2001) What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental & Resource Economics* 20(3): 241–254.

Lo, A.Y. and Spash, C.L. (2013) Deliberative monetary valuation: in search of a democratic and value plural approach to environmental policy. *Journal of Economic Surveys* 27(4): 768–789.

Loomis, J. (2011) What’s to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys* 25(2): 363–370.

Lupia, A. (1994) Shortcuts versus encyclopedias – information and voting behavior in California insurance reform elections. *American Political Science Review* 88(1): 63–76.
Lupia, A. and Matsusaka, J.G. (2004) Direct democracy: new approaches to old questions. *Annual Review of Political Science* 7: 463–482.

Lupia, A. and McCubbins, M.D. (1998) The democratic dilemma: can citizens learn what they need to know? New York: Cambridge University Press.

Maital, S. (1979) Measurement of net benefits from public goods: a new approach using survey data. *Public Finance Finances Publiques* 34(1): 85–99.

McFadden, D. (1994) Contingent valuation and social choice. *American Journal of Agricultural Economics* 76(4): 689–708.

McFadden, D. (1999) Rationality for economists? *Journal of Risk and Uncertainty* 19(1-3): 73–105.

McFadden, D. (2001) Economic choices. *American Economic Review* 91(3): 351–378.

Mitchell, R.C. and Carson, R.T. (1989) Using surveys to value public goods: the contingent valuation method. Washington, D.C.: Resources for the Future.

Murphy, J.J., Allen, P.G., Stevens, T.H. and Weatherhead, D. (2005) A meta-analysis of hypothetical bias in stated preference valuation. *Environmental & Resource Economics* 30(3): 313–325.

Noam, E.M. (1982) Demand functions and the valuation of public goods. *Public Choice* 38(3): 271–280.

Pendse, D. and Wyckoff, J.B. (1976) Measurement of environmental trade-offs and public policy: a case study. *Water Resources Research* 12: 919–930.

Pommerehne, W.W. (1987) Präferenzen für öffentliche Güter. Ansätze zu ihrer Erfassung. J.C.B Mohr, Tübingen.

Preston, I. and Ridge, M. (1995) Demand for local public spending: evidence from the British Social Attitudes Survey. *The Economic Journal* 105(430): 644–660.

Randall, A., Ives, B. and Eastman, C. (1974) Bidding games for valuation of aesthetic environmental improvements. *Journal of Environmental Economics and Management* 1(2): 132–149.

Rheinberger, C. (2011) A mixed logit approach to study preferences for safety on Alpine roads. *Environmental & Resource Economics* 49(1): 121–146.

Rongen, G. (1995) Efficiency in the provision of local public goods in Norway. *European Journal of Political Economy* 11: 235–264.

Rubinfield, D.L. (1977) Voting in a local school election – a micro-analysis. *Review of Economics and Statistics* 59(1): 30–42.

Rubinfield, D.L. and Shapiro, P. (1989) Micro-estimation of the demand for schooling: evidence from Michigan and Massachusetts. *Regional Science and Urban Economics* 19: 381–398.

Rubinfield, D.L., Shapiro, P. and Roberts, J. (1987) Tiebout bias and the demand for local public schooling. *Review of Economics and Statistics* 69(3): 426–437.

Satterthwaite, M.A. (1975) Strategy-proofness and Arrow conditions – existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory* 10(2): 187–217.

Schläpfer, F. (2011) Access to party positions and preference formation: a field experiment. *Swiss Political Science Review* 17(1): 75–91.

Schläpfer, F. and Hanley, N. (2006) Contingent valuation and collective choice. *Kyklos* 59: 115–135.

Schläpfer, F., Roschewitz, A. and Hanley, N. (2004) Validation of stated preferences for public goods: a comparison of contingent valuation survey response and voting behaviour. *Ecological Economics* 51(1-2): 1–16.

Schläpfer, F. and Schmitt, M. (2007) Anchors, endorsements, and preferences: a field experiment. *Resource and Energy Economics* 29(3): 229–243.

Schläpfer, F., Schmitt, M. and Roschewitz, A. (2008) Competitive politics, simplified heuristics, and preferences for public goods. *Ecological Economics* 65(3): 574–589.

Schokkaert, E. (1987) Preferences and demand for local public spending. *Journal of Public Economics* 34: 175–188.

Shabman, L. and Stephenson, K. (1996) Searching for the correct benefit estimate: empirical evidence for an alternative perspective. *Land Economics* 72(4): 433–449.
Shapiro, P. and Deacon, R.T. (1996) Estimating the demand for public goods: comments and extensions. In D. J. Bjornstad and J. R. Kahn (eds.), The Contingent Valuation of Environmental Resources. Methodological Issues and Research Needs (pp. 244–262). Cheltenham, U.K.: Edward Elgar

Smith, V.L. (2003) Constructivist and ecological rationality in economics. American Economic Review 93: 465–508

Strauss, R.P. and Hughes, G.D. (1976) A new approach to the demand for public goods. Journal of Public Economics 6(3): 191–204.

Sunstein, C.R. and Thaler, R.H. (2003) Libertarian paternalism is not an oxymoron. University of Chicago Law Review 70(4): 1159–1202.

Sunstein, C.R. (2014) The real world of cost-benefit analysis: thirty-six questions (and almost as many answers). Columbia Law Review 114: 167–2012.

Tversky, A. and Kahneman, D. (1981) The framing of decisions and the psychology of choice. Science 211(4481): 453–458.

USEPA (2009) Valuing the protection of ecological systems and services. A report of the EPA Science Advisory Board. Washington, D.C.: United States Environmental Protection Agency.

Vossler, C.A. and Kerkvliet, J. (2003) A criterion validity test of the contingent valuation method: comparing hypothetical and actual voting behavior for a public referendum. Journal of Environmental Economics and Management 45(3): 631–649.

Vossler, C.A., Kerkvliet, J., Polasky, S. and Gainutdinova, O. (2003) Externally validating contingent valuation: an open-space survey and referendum in Corvallis, Oregon. Journal of Economic Behavior & Organization 51(2): 261–277.