# Public Resource Allocation, Strategic Behavior, and Status Quo Bias in Choice Experiments


## Abstract

Choice experiments, a survey methodology in which consumers face a series of choice tasks requiring them to indicate their most preferred option from a choice set containing two or more options,  are used to generate estimates of consumer preferences to determine the appropriate allocation of public resources to competing projects or programs.  The analysis of choice experiment data typically relies on the assumptions that choices of the non-status quo option are demand-revealing and choices of the status quo option are not demand-revealing, but rather, reflect an underlying behavioral bias in favor of the status quo.  This paper reports the results of an experiment which demonstrates that both of these assumptions are likely to be invalid.  We demonstrates that choice experiments for a public good are vulnerable to the same types of strategic voting that affect other types of multiple-choice voting mechanisms.  We show that due to the mathematics of choice set design, what is actually strategic voting is often interpreted as a behavioral bias for the status quo option.  Therefore, we caution against using current choice experiment methodologies to inform policy making about public goods.



Keywords: choice experiment, strategic voting, status quo bias, public goods experiment

JEL codes: H41, C91, C92

**Public Resource Allocation, Strategic Behavior, and Status Quo Bias in Choice Experiments**

## 1. Introduction

This paper cautions against the increasing use and dominance of the choice experiment method to generate estimates of the benefits from public policies and programs. "Choice experiments (CEs)," a survey method used to estimate consumer preferences, were initially developed and are widely used in the context of private goods for the fields of marketing and transport to study consumer preferences and predict market shares for new products (e.g. Louviere and Woodworth, 1983; Louviere, 1984, 1988; Revelt and Train, 1998; Brownstone and Train, 1999; McFadden, 2001; Hensher, Rose and Greene, 2015). However, the methodology was soon applied to informing public programs and policies in the environment and health domains, whose outputs are largely, or solely, public goods (e.g. Australian Energy Market Operator, 2014; Cameron and DeShazo, 2013; Emmerson and Metcalfe, 2013; Queen's University Belfast and Perceptive Insight, 2015). Choice experiment mechanisms can take a variety of forms such as one-shot or repeated choice tasks. On each choice task, participants are presented with a choice set containing two or more choice options from which they must choose one option. In order to be consistent with market settings in which consumers always have the option not to purchase, in most choice experiments, one of the options in each choice task will always be a 'no purchase' (in CEs for private goods) or 'status quo' (in CEs for public goods) option. Figure 1 reports example choice tasks from choice experiment surveys used in three very different settings: to estimate the demand for quality-differentiated beef (Lusk and Schroeder, 2004); to determine the public's preferences for the use of public lands for endangered species protection vs. military training (Smith and McKee, 2007); and to estimate

Sample choice task for quality-differentiated beef steaks (Lusk and Schroeder, 2004, p. 469):

| Scenario 11 | Steaks | | | | | |
|---|---|---|---|---|---|---|
| | Generic $6.75 | Guaranteed Tender $7.88 | Natural $9.00 | USDA Choice $5.63 | Certified Angus Beef $7.88 | None of These |
| I would choose . . . | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

Sample choice task for use of public land for endangered species protection vs. military training (Smith and McKee, 2007, p. 232):

**Question 5**

**Suppose the following three options were the only options available for managing the balance between endangered species encroachment and the DOD training mission. mission. Please indicate which option you prefer by placing an "X" in one of the boxes below.**

| Feature | Option A | Option B | Option C: Current Status |
|---|---|---|---|
| Payment per household | $50 | $75 | None |
| Impact on species survival | High | Medium | Low |
| Impact on soldiersí readiness | Medium | Low | High |

☐ I would choose A

☐ I would choose B

☐ I would choose C

Sample choice task for a cervical cancer screening program (Ryan and Wordsworth, 2000, p. 511)

| • **Choice 1** | Option A | Option B | |
|---|---|---|---|
| Time between smears (years) | 3 | 1 | |
| Time for results (days) | 28 | 10 | |
| Chance of being recalled | 11% | 20% | |
| Chance of abnormality | 10% | 10% | |
| Chance of dying from cervical cancer | 2% | 0.8% | |
| Cost of each smear (£) | 5 | 20 | |
| | Prefer Option A | Prefer Option B | Prefer no screening |
| Which Option would you prefer? | ☐ | ☐ | ☐ |

Figure 1. Sample Choice Tasks from Choice Experiment Surveys

preferences for a cervical cancer screening program (Ryan and Wordsworth, 2000). The common features of these choice tasks are (1) Choice options which are described in terms of their characteristics, termed attributes, which vary across options; (2) The inclusion of more than two options from which respondents may choose; (3) The inclusion of a price or cost attribute for each option (which enables researchers to estimate willingness to pay); and (4) The inclusion of a no purchase/no policy change (status quo) option. Due to the gains in statistical efficiency of preference estimates that result from asking respondents to choose from more than two options and presenting respondents with more than one choice task (both of which generate more information about each respondent's preferences, resulting in greater statistical efficiency), most choice experiment surveys employ the repeated multiple choice mechanism, in which participants complete a series of choice tasks similar to those presented in Figure 1. As respondents proceed through the tasks, the characteristics of the purchase/non-status-quo options change, while the characteristics of the no purchase/status quo option remain the same. Using logit-based econometric techniques (McFadden, 1986, 2001), researchers can use respondents' stated choices to estimate respondents' willingness to pay for a good or policy as well as their marginal willingness to pay for the attributes of a good or policy.

Although we might expect to obtain reliable preference estimates in the private goods domain, the application of the choice experiment domain to the public goods environment is potentially problematic. In particular, the possibility of strategically biased preference estimates in the public goods environment has been neglected in the literature to date. One of the most common biases in choice experiments that researchers have found is apparent over-choice of the no purchase/status quo option (e.g. Collins and Vossler, 2009; Adamowicz et al., 2011), termed status quo bias. This apparent bias in favor of the status quo option can significantly reduce the estimated value of the public good (Adamowicz et al., 1998), potentially resulting in misallocations of public resources. Although there exist multiple explanations for the source of this bias in the literature,[1] as well as econometric methods (Adamowicz et al., 1998) to adjust benefit estimates for the presence of status quo bias, none of these studies considers the role that the underlying incentives of the choice mechanism in a public goods environment may play in generating bias. This may be due to the inherent difficulties involved in identifying strategic behavior in field survey data.

This paper takes the choice experiment out of the field and into the laboratory in order to establish whether choice experiments can generate unbiased estimates of the benefits from public goods. By taking this approach, we are able to demonstrate unambiguously that choice experiments for public goods are vulnerable to the same type of strategic voting that occurs in multi-candidate elections. Strategic choices fall on the second-best option which, because of the mathematics of combinatorial choice set design[2], is the status quo option in a disproportionate number of choice tasks. The experimental results demonstrate that it is inappropriate to interpret the results from choice experiments as either fully demand-revealing, or as demand-revealing with an adjustment for status quo bias. Both interpretations of choice experiment data can result in biased estimates of preferences for public goods, and thus misallocations of public resources. Because of this potential problem, we argue that current choice experiment methods to value public goods should be used with caution.

[1] See, for example: List et al. (2006); Taylor et al. (2010); Carlsson et al. (2007); Bateman et al. (2008); Collins and Vossler, 2009; Day and Pinto-Prades, 2010; Day et al. (2012), and Aravena et al. (2014).

[2] By the "mathematics of combinatorial choice set design," we mean the method by which individual choice options with different levels of attributes are combined into groups of options, termed choice sets. During a choice experiment survey, respondents are presented with a series of choice tasks, in which they are asked to choose one option out of each choice set. Their choice is usually interpreted to indicate which option in the set they most prefer.

We provide three forms of support for this argument: (1) Well-known results from the theoretical and applied literature on strategic voting demonstrating voters' incentives to vote for their second favorite option in races involving three or more candidates; (2) A mathematical model that demonstrates how the mathematics of combinatorial choice set design can result in choice experiments in which a large fraction of choice tasks contain the status quo option as the second-best choice; and (3) Evidence from an experiment based on three fractional factorial designs that demonstrates that behavior in such choice experiments is consistent with the predictions of voting theory. The remainder of this paper lays out the argument by presenting the theoretical and mathematical models in section 2, the experimental design in section 3, the results in section 4, and concluding remarks in section 5.

## 2. Theory and Simulations

### *2.1 Voting in Multi-Candidate Elections*

As the theory of voting is well-known, we provide only a brief summary of it here. Farquharson (1969) provides a comprehensive overview of the topic. Consider a group of $n$ agents voting on the choice of a public program or policy. Suppose that agents choose from three discrete public goods, $g \in \{A, B, C\}$. Each agent $i$ has a complete, reflexive, and transitive preference ordering over the three proposed programs, which can be described by a utility function $u_i(g)$. In order to determine which program to provide, the public agency holds a vote. Each agent $i$ casts a vote $v_i(g)$. Under a plurality voting rule, the agency's decision rule is to provide the public program which receives the most votes, $g^*$. Thus, each agent's problem is to choose their vote $v_i(g)$ to maximize their utility, $u_i(g)$, conditional on the distribution of votes of the other group members. In the case of uniform priors, in which agent $i$ believes that the other group members' votes are distributed equally across the three outcomes, agent $i$'s best response is to vote for his or her most preferred option. In this case, agent *i*'s vote is demand revealing. However, if agent *i* has non-uniform priors about the distribution of other votes, they may have the incentive to vote for an option other than their most preferred option. For example, if agent *i*'s preferences are such that $u_i(A) > u_i(B) > u_i(C)$, but their priors are that:

$$\sum_{j \neq i} v_j(C) > \sum_{j \neq i} v_j(B) > \sum_{j \neq i} v_j(A) \tag{1}$$

then agent *i* can potentially improve their utility by voting for their second most-preferred option rather than their most preferred option, in order to prevent their least-preferred option from winning. In such cases, agent *i*'s vote is not demand revealing.

Strategic voting of this type is perhaps the most well-known result of voting theory. It has been demonstrated to occur in laboratory settings under conditions of full information about the distribution of the preferences of the other voters (Felsenthal et al., 1988; Forsythe et al., 1993) and in field settings in which voters have only partial information about the distribution of other voters' preferences (Fujiwara, 2011; Kawai and Watanabe, 2013). Given that the incentive structure of each choice task in a choice experiment mechanism for a public good is identical to that of a multi-candidate election, it is reasonable to expect that a similar type of strategic voting might take place in choice experiment surveys. We demonstrate below that due to the mathematics of combinatorial choice set design, strategic votes for the second best option are likely to be disproportionately for the status quo option.

*2.2. Combinatorial Choice Set Design and the Second-Best Option*

We begin with a few definitions:

Definition 1: A *choice experiment* is a survey in which a respondent faces a series of choice tasks such as the ones illustrated in Figure 1. On each choice task, the respondent faces a choice set from which s/he must select one option.

Definition 2: A *choice set* is a group of two or more *choice options*. The number of options in the choice sets illustrated in Figure 1 ranges from three (for the land use and health studies) to six (for the beef study) options. In most choice experiment surveys, including the ones illustrated in Figure 1, one of the choice options is typically the no purchase/status quo option.

Definition 3: A *choice option* is a version of the good that the respondent may choose. Each choice option is described in terms of the amount of each *attribute* that it possesses. Choice options in Figure 1 include five different types of steak and a no purchase option, two alternative land use policies along with a keep current policy (status quo) option, and two alternative cancer screening program options along with a no screening option.

Definition 4: An *attribute* is a characteristic of a good. The amount of an attribute that a good possesses may be described using discrete or continuous numerical values. The cost or price of an option is usually an attribute in order to facilitate the estimation of a respondent's willingness to pay. In Figure 1, the attributes of the steaks are limited to their price and a descriptor of their quality (e.g. Guaranteed Tender). The attributes of the land use policy include the cost of each policy per household and the impact of the policy on species survival and soldiers' readiness (both of which can take on levels of low, medium, or high). Attributes of the cancer screening program include the time between smears, the time to obtain results, chance of being recalled, chance of an abnormality being discovered, chance of dying from cervical cancer, and the cost.

In the derivation that follows, we consider fundamental principles of attribute-balanced choice set design applied to choice experiments for public goods. An attribute balanced choice set is a choice set in which every choice option is described using the same number of attributes, and each attribute has the same number of possible levels. In Figure 1, the choice sets in the land use study are attribute-balanced. The mathematical results below easily generalize to choice sets in which different attributes have different numbers of possible levels.[3] In choice sets in which each of $m$ attributes has $n$ levels, there are $n^m$ possible choice options which can be combined into choice sets. One of more of these options are combined with the status quo/no purchase option (whose attributes are fixed) to create a choice set. To illustrate the effect of this choice set design process, we let $\pi$ represent the proportion of the non-status-quo choice options that are preferred by the voting public to the status quo. We make the simplifying assumption that all voters have the same value of $\pi$. The result generalizes to the more realistic assumption that there is a distribution of values of $\pi$ across the voting population. Therefore, given the values of $n$, $m$, and $\pi$, $\pi n^m$ choice options are preferred to the status quo option, and $(1 - \pi)n^m$ choice options are not preferred to the status quo.

A choice experiment in which respondents face a sufficient number of choice tasks that they see all the possible combinations of choice options in choice sets of a given size is termed a full factorial choice experiment. We consider a full factorial choice experiment in which on each choice task, respondents must choose from a three-option choice set consisting of two options plus the status quo. We focus our analysis on this type of choice experiment both because it is the simplest choice experiment design in which the choice tasks contain more than two options, and because choice experiment designs of this type are commonly used in applied public goods settings (Ferrini and Scarpa, 2007). In a full factorial choice experiment in which each choice set contains two options plus the status quo, there are $n^m(n^m - 1)$ total possible choice sets. In $(\pi n^m)(\pi n^m - 1)$ of these choice sets, the status quo will be the least-preferred option. In $[(1 - \pi)n^m][(1 - \pi)n^m - 1]$ of these choice sets, the status quo will be the most-preferred option. This leaves:

[3] A proof of this generalized result is available from the authors.

| Number of Attribute Levels ($n$) | | Number of Attributes ($m$): 2 | 3 | 4 |
|---|---|---|---|---|
| | 2 | 0.833 | 0.786 | 0.767 |
| | 3 | 0.781 | 0.760 | 0.753 |
| | 4 | 0.767 | 0.754 | 0.751 |

Table 1. Proportion of choice sets with status quo as first-best or second-best option when $\pi = 0.5$

$$n^m(n^m - 1) - \{ \pi n^m(\pi n^m - 1) + [(1 - \pi)n^m][(1 - \pi)n^m - 1]\} \qquad (2)$$

choice sets in the full factorial design in which the status quo option is the second-best option. In the limit (as either $n \to \infty$ *or* $m \to \infty$), the ratio of this number to the total number of choice sets is $2\pi(1 - \pi)$. Furthermore, the limiting fraction of total possible choice sets in the full factorial design in which the status quo is either the most-preferred or second-most-preferred option is $(1 - \pi^2)$.

*2.3. Calculations and Simulation Results*

This derivation demonstrates that the status quo is likely to land in the first-best or second-best position in a large fraction of choice sets. In a full factorial choice experiment in which half the options are preferred to the status quo and half are not ($\pi = 0.50$), the status quo will reside in the first-best or second-best position in $(1 - (0.5)^2) =$ 75% of the choice sets. This result is for large values of $n$ or $m$ and all possible choice sets. Table 1 reports the proportion of choice sets in which the status quo option is first-best or second-best for small values of $n$ or $m$ when $\pi = 0.5$. The results in Table 1 demonstrate that the limiting case is reached very quickly.

| Option 1 | Option 2 | Option 3 (current portfolio) |
|---|---|---|
| 1 unit of **A** | 2 units of **A** | 1 unit of **A** |
| 3 units of **B** | 1 unit of **B** | 1 unit of **B** |
| **Cost** = 5 | **Cost** = 10 | **Cost** = 0 |

Figure 2. Sample Choice Set from Simulation Exercise

Typically, field choice experiments do not employ full factorial designs, as these would require respondents to complete an inordinate number of choice tasks. Choice experiments which use a subset of the choice sets from the full factorial design are termed fractional factorial choice experiments. We employ three techniques to derive fractional factorial choice experiment designs from the full factorial to examine the extent to which the predictions of the mathematical model for the full factorial design hold for fractional factorial designs. We derive designs for choice experiments containing 12 choice tasks in which the choice sets are either utility balanced (UBAL) (Huber and Zwerina, 1996), orthogonal on the attribute level differences (OOD, Kanninen, 2002; Street et al., 2005), or randomly drawn (RAND) from the full factorial design. Each option in each choice set has three attributes, one of which is the cost, and each attribute has two possible levels. The status quo option contains baseline levels of each attribute at no cost. Figure 2 illustrates a sample choice set used in the simulation. In order to generate a simulated voter's rank-ordering of the options in a choice set, we assume the following utility function:

Utility = (Marginal value of attribute A × Units of attribute A) + (Marginal value of attribute B × Units of attribute B) – Cost

By varying a voter's marginal values of the attributes in the equation above, we can vary the value of $\pi$ and calculate a voter's simulated utility for each choice option, as well as their rank-ordering of each choice option in each choice set. Table 2 reports the results of the simulations, and demonstrates that the mathematical predictions based on the value of $\pi$ applied to a full

| | Fraction of Choice Sets in Which the Status Quo is Second-Best | | | Fraction of Choice Sets in Which the Status Quo is First- or Second-Best | | |
|---|---|---|---|---|---|---|
| | $\pi = 0.125$ | $\pi = 0.50$ | $\pi = 0.75$ | $\pi = 0.125$ | $\pi = 0.50$ | $\pi = 0.75$ |
| Full Factorial Prediction: $n = 2$, $m = 3$ | 0.25 | 0.57 | 0.43 | 1.00 | 0.79 | 0.46 |
| Full Factorial Prediction: Limiting Case | 0.22 | 0.50 | 0.38 | 0.98 | 0.75 | 0.44 |
| OOD Fractional Factorial Choice Experiment Design | 0.25 | 1.00 | 0.50 | 1.00 | 1.00 | 0.50 |
| UBAL Fractional Factorial Choice Experiment Design | 0.08 | 0.33 | 0.33 | 1.00 | 0.75 | 0.42 |
| RAND Fractional Factorial Choice Experiment Design | 0.17 | 0.67 | 0.42 | 1.00 | 0.75 | 0.50 |

Table 2. Full-Factorial Predictions of Fractions of Choice Sets Containing Status Quo as Second-Best or First- or Second-Best Option vs. Actual Results for Fractional Factorial Choice Experiment Designs

factorial design are reasonable estimates of the fraction of choice sets in a fractional factorial design that will contain the status quo as the second-best or first-best or second-best option.

Although the status quo option is likely to land in the second-best position in a large percentage of choice sets, voters only have the incentive to cast non-demand revealing votes for the second-best option if they have non-uniform priors about the distribution of other votes. In order to examine whether this second condition is satisfied, we employ the orthogonal on the attribute level differences (OOD) choice experiment design from the simulations above, and ask what would be the distribution of preferences over choice options in each choice set if the distribution of $\pi$ were uniform. Table 3 reports the results of this exercise. Uniform distributions of voter preferences over choice options are clearly not the norm, which means that it may not be reasonable to assume that voters have uniform priors about the distribution of votes. Thus, the conditions necessary for strategic voting appear to be present in choice experiments, and as a result of the mathematics of combinatorial choice set design, many of the strategic votes are

| Choice Set | Share who Prefer Option 1 | Share Who Prefer Option 2 | Share Who Prefer Status Quo |
|---|---|---|---|
| 1 | 0% | 57% | 43% |
| 2 | 86% | 0% | 14% |
| 3 | 86% | 0% | 14% |
| 4 | 0% | 57% | 43% |
| 5 | 71% | 0% | 29% |
| 6 | 14% | 29% | 57% |
| 7 | 0% | 86% | 14% |
| 8 | 29% | 14% | 57% |
| 9 | 0% | 71% | 29% |
| 10 | 29% | 14% | 57% |
| 11 | 0% | 71% | 29% |
| 12 | 57% | 0% | 43% |

Table 3. Distribution of voter preferences assuming a uniform distribution of $\pi$

likely to be for the status quo option, which will frequently fall in the second best position in a choice set.

## 3. Experimental Design

### *3.1 Rules of the Game*

The experiment is designed to simulate a field choice experiment. In the experiment, subjects in groups of nine, which remain the same for the entire experimental session, vote on a series of twelve ballots for a public good, which for the purposes of the experiment is characterized as a group investment for which earnings vary across subjects. Each ballot contains a choice set containing three options. Each option has two attributes and a cost. There are two possible levels of each attribute and the cost. We reserve additional details about the choice sets for the following section on induced values.

The order of ballots is randomized across subjects. Subjects complete each ballot one by one, and submit them to the experimental moderator. All ballots are cast privately, and each subject only sees the information on his or her own ballot. After the twelve rounds of voting are completed, one ballot is randomly selected from among the twelve to be binding using the roll of a 12-sided die. The votes on the binding ballot are counted and a plurality voting rule is applied, so that the choice option on the binding ballot with the most votes wins. Two-way ties are resolved using the toss of a fair coin, and three-way ties are resolved via the roll of a fair six-sided die. The winning option is announced, and subjects compute their earnings for the winning option using the formula:

$$\text{Earnings} = 6 + (\text{Marginal value of attribute A} \times \text{Units of attribute A}) + (\text{Marginal value of attribute B} \times \text{Units of attribute B}) - \text{Cost}$$

Subjects know all of the rules of the game at the beginning of the experiment. Each subject knows that other members of the group may face a different set of choice options on any given ballot and that other subjects may also have different marginal values for each attribute, but no subject knows the overall distribution of marginal values or possible choice options. Figure 3 presents a sample ballot.

**BALLOT**

**Your value of A is: 3**
**Your value of B is: 7**

| Option 1 | Option 2 | Option 3 (current portfolio) |
|---|---|---|
| 1 unit of **A** | 2 units of **A** | 1 unit of **A** |
| 3 units of **B** | 1 unit of **B** | 1 unit of **B** |
| **Cost** = 5 | **Cost** = 10 | **Cost** = 0 |

Please select which option you wish to vote for below:

☐ Option 1

☐ Option 2

☐ Option 3

Figure 3. Sample Experimental Ballot

*3.2. Choice Set Designs*

Each choice option contains three attributes (A, B, and Cost) with two levels. Attribute A has possible levels of 1 or 2, Attribute B has possible levels of 1 or 3, and the Cost is either 5 or 10. This design is the simplest possible design that mimics a field choice experiment. A full factorial choice experiment design contains $2^3 = 8$ possible choice options which can be combined into 56 possible multiple choice sets containing two different options and the status quo option. The status quo option contains 1 unit of attribute A, 1 unit of attribute B, and is available at zero cost.

In order to test whether behavior is different in different types of choice experiment designs, we employ three methods to reduce the number of choice sets from the full factorial: orthogonality on the attribute level differences (OOD) (Kanninen, 2002; Street et al., 2005), utility balance (UBAL) (Huber and Zwerina, 1996), and random (RAND). In an OOD choice set in which each attribute has only two levels, if one attribute of one option in the set has the high level, that same attribute in the other option in the choice set will be assigned the low level. Thus, the smallest possible choice experiment that results from a $2^3$ factorial design is one containing 12 choice sets. For consistency, choice experiments designed using the random and utility balanced methodologies also contain 12 choice sets. Normally, to create utility-balanced choice sets, the designer would have to make assumptions regarding which choice options are likely to be considered better or worse by respondents in order to generate the property of utility balance. In an induced-value setting, subjects' induced utility for each choice option is known via the earnings equation above. Thus, we create a different set of utility-balanced choice sets for each set of induced values. All choice set designs are reported in the supplemental materials.

*3.3. Induced Values*

The experimental design incorporates three sets of induced values, corresponding to values of $\pi$ of 0.125, 0.50, and 0.75. Table 4 reports the marginal values of attributes A and B associated with each value of $\pi$.

The combination of three methods to create fractional factorial choice experiment designs and three sets of induced values results in nine experimental conditions. Within an experimental session, which could contain up to 27 participants, subjects are randomly assigned to a group of nine that play either the OOD, UBAL, or RAND version of the choice experiment.[4] Within a group of nine, each subject is randomly assigned an induced value. Subjects are blinded to the method by which their choice sets were created, the induced values of the other individuals in

[4] If the number of subjects who showed up was not divisible by nine, the unassigned subjects were invited to participate in a different experimental session at a later time. Prior to starting the experiment, all subjects completed an informed consent process. Subjects were free to leave at any time.

| $\pi$ | Marginal value of attribute A | Marginal value of attribute B |
|---|---|---|
| 0.125 | 3 | 2 |
| 0.50 | 3 | 7 |
| 0.75 | 11 | 7 |

Table 4. Induced Values

their group, and the overall distribution of induced values. Moderators are also blinded to the distribution of induced values in the group. Two hundred and seventy student volunteers from a U.S. university participated in the experiment, resulting in a sample of 30 subjects for each choice experiment design-induced value combination. Since each subject votes 12 times, this results in a total of 360 observations per choice experiment design-induced value combination, or 1080 total observations for each choice experiment design methodology. The experiment takes 40-45 minutes to complete, and average experimental earnings are $15-$20.[5]

**4. Results and Discussion**

The results to be reported demonstrate that (1) Strategic voting occurs in choice experiments for public goods; (2) Strategic votes for the status quo option happen a predictable percentage of the time; and (3) Most choices of the status quo option are not examples of status quo bias, but rather, are demand-revealing. As a consequence, it is inappropriate to interpret the data from choice experiments as either fully demand revealing, or as subject to a behavioral bias in favor of the status quo. Public policy choices informed by analyses based on either of these two assumptions may result in inefficient allocations of public resources. In the discussion that

[5] The experiment was conducted under the oversight of the university's Institutional Review Board (IRB). All experimental instructions are available in the online supplementary materials. Experimental data are available from the corresponding author upon request.

| | Mean | Robust Std. Err. | 95% Conf. Int. |
|---|---|---|---|
| By Method to Create Fractional Factorial Choice Set Designs | | | |
| OOD | 0.79 | 0.051 | [0.69, 0.89] |
| UBAL | 0.82 | 0.050 | [0.72, 0.92] |
| RAND | 0.83 | 0.046 | [0.74, 0.92] |
| By Fraction of Choice Options Preferred to the Status Quo ($\pi$) | | | |
| $\pi = 0.125$ | 0.87 | 0.035 | [0.80, 0.94] |
| $\pi = 0.50$ | 0.79 | 0.048 | [0.70, 0.89] |
| $\pi = 0.75$ | 0.73 | 0.064 | [0.61, 0.86] |

Table 5. Fraction of Non-Demand Revealing Votes for the Second-Best Alternative

follows, all p-values are calculated using standard errors clustered at the individual subject level to account for within-subject correlation of errors.

*4.1 Result 1: Strategic voting occurs in choice experiments for public goods.*

One hundred fourteen subjects (42% of all subjects) cast at least one vote for the second-best option in a choice set during the experiment. As reported in Table 5, votes for the second-best option constitute over 70% of all non-demand revealing votes in the experiment. Conditional on a vote being non-demand revealing, there are no significant differences in the fractions of votes for the second-best alternative across choice experiment design treatments ($0.55 \leq p \leq 0.90$). Subjects with low values of $\pi$ are significantly more likely than other subjects to vote for the second-best option ($p < 0.01$). There are no other significant differences in the nature of non-demand revealing behavior across subject types.

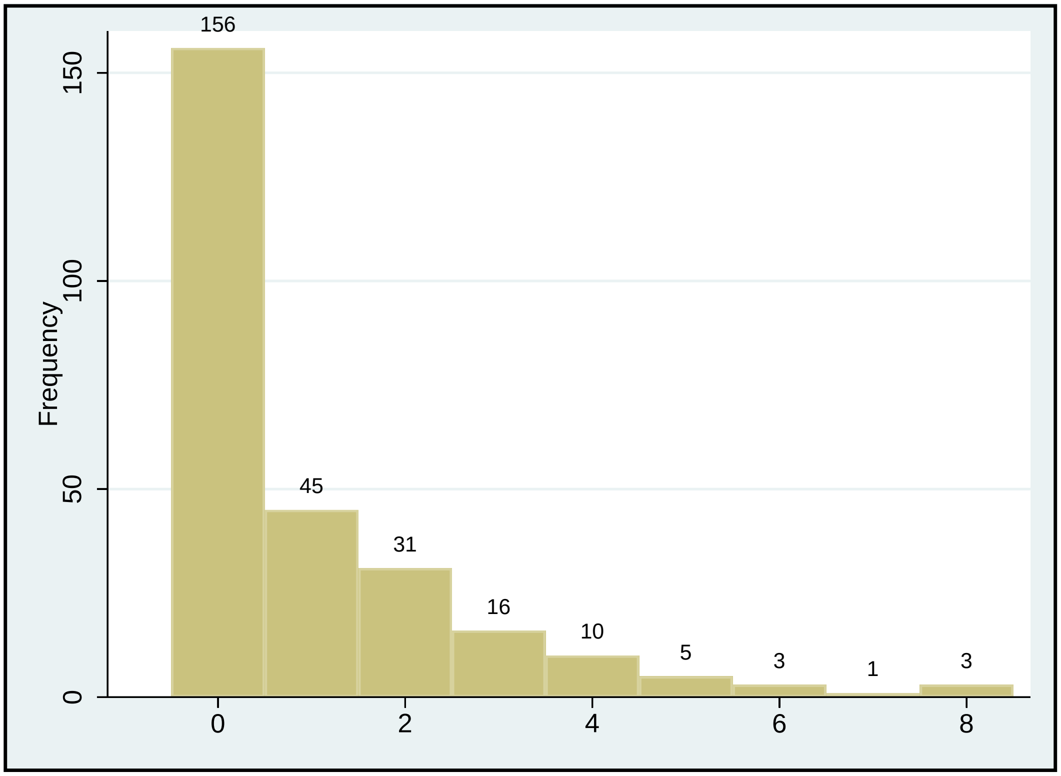


Figure 4. Frequency of number of second-best votes per subject

Figure 4 reports the frequency of the total number of second-best votes cast by individual subjects. The vast majority of subjects who cast second-best votes cast one or two votes for the second-best option out of a possible 12. Subjects tend to vote for the second-best option when there is a very bad third option in a choice set. Including the status quo option, there are nine possible choice options which can be combined into choice sets. Figure 5 reports the ranking out of 9 of the worst option in each choice set for those ballots for which non-demand revealing votes for the second-best alternative were cast. When subjects cast a non-demand revealing vote for the second-best alternative, 65% of the time the ballot contained a third option which ranked

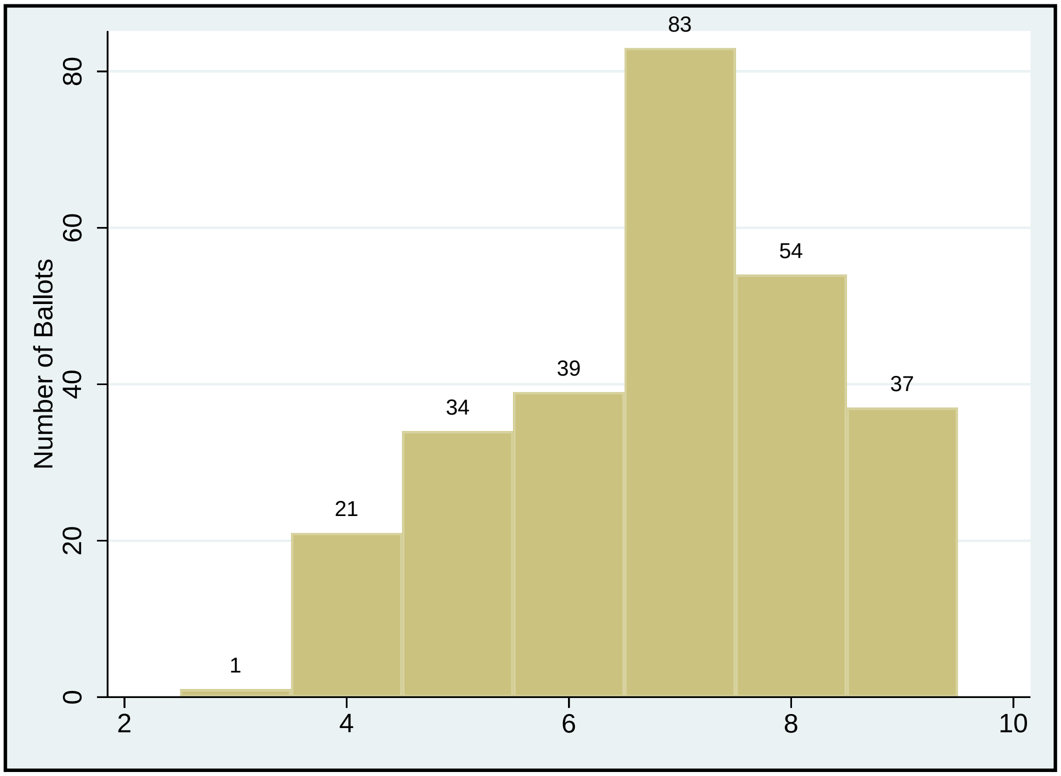


Figure 5.  Ranking (out of 9) of the worst option in a choice set when a non-demand revealing vote for the second best option was cast

7th, 8th, or 9th out of the possible nine alternatives.  Seventy-nine percent of these ballots contained a third choice alternative which fell in the bottom half of the ranking.  The result that most non-demand revealing votes are for the second-best option combined with the result that these votes tend to occur when there is a bad third option in the choice set, and tend to be cast more often by the $\pi = 0.125$ subjects who have many more bad options in their choice sets demonstrate that non-demand revealing behavior is not random, but rather follows a pattern consistent with a model of strategic voting.

The overall rate of strategic voting in the experiment is 8%. This rate of strategic voting is well within the range of rates of strategic voting previously observed in the laboratory (as high as 50%, as reported in Felsenthal et al., 1988, and Forsythe et al., 1993) and in the field (1.2%, as reported in Kawai and Watanabe, 2013). In most laboratory studies of strategic voting, subjects have complete information about the distribution of other voters' preferences, giving full information about non-uniform priors and making the incentives for strategic voting completely transparent. Obviously, such information is not available to voters in the field. In this experiment, subjects have complete information about their own preferences (as in other lab studies), but must form their own priors about the distribution of preferences of other members of the group (as in field studies). Given that this experimental design contains elements of both the lab and the field, it is not surprising to observe rates of strategic voting that are consistent with and between previous lab and field results.

*4.2. Result 2: Strategic votes for the status quo option happen a predictable percentage of the time.*

The first five rows of Table 6 report the predicted fractions of votes for the second-best option that are the status quo option. The last three rows of Table 6 show the observed fraction of second best votes which are for the status quo. By comparing the last three rows to the rows above them, it is possible to examine the extent to which strategic votes for the status quo are consistent with the mathematical predictions based on the value of $\pi$.

| | Fraction of Choice Sets in Which the Status Quo is Second-Best | | |
|---|---|---|---|
| | $\pi = 0.125$ | $\pi = 0.50$ | $\pi = 0.75$ |
| Full Factorial Prediction: $n = 2$, $m = 3$ | 0.25 | 0.57 | 0.43 |
| Full Factorial Prediction: Limiting Case | 0.22 | 0.50 | 0.38 |
| OOD Fractional Experimental Design | 0.25 | 1.00 | 0.50 |
| UBAL Fractional Experimental Design | 0.08 | 0.33 | 0.33 |
| RAND Fractional Experimental Design | 0.17 | 0.67 | 0.42 |
| Experimentally Observed Fraction of Second-Best Votes for the Status Quo | | | |
| Observed: OOD Experimental Results | 0.03 | 1.00 | 0.40 |
| (Robust SE) | (0.04) | (N/A) | (0.15) |
| Observed: UBAL Experimental Results | 0.04 | 0.07 | 0.47 |
| (Robust SE) | (0.03) | (0.03) | (0.18) |
| Observed: RAND Experimental Results | 0.02 | 0.80 | 0.39 |
| (Robust SE) | (0.02) | (0.16) | (0.10) |

Table 6. Predicted and Observed Votes for Second-Best Options that are the Status Quo Option

Subjects may deviate from these predictions by casting a second-best vote for the status quo more often than predicted by the model, or less often, depending on when they choose to cast votes for a second-best option. If they are casting second-best votes more often in choice tasks in which the status quo is not in the second-best position, the observed fraction of second-best votes for the status quo option will be less than predicted by the value of $\pi$. If they are casting second-best votes more often on choice occasions where the status quo is in the second-best position, then the observed fraction of second-best votes for the status quo will be greater than predicted by the value of $\pi$. Both outcomes are observable in Table 6. The choice tasks on which subjects may choose to cast second-best votes depends on a subject's priors and how bad the third alternative is for them, which varies across choice sets and by value of $\pi$. The exceptions to this are the OOD choice sets for $\pi = 0.50$ subjects. Since all of the OOD choice sets contain the status quo in the second-best position, all second-best votes are necessarily for the status quo alternative. The key message of Table 6 is, with the exception of the UBAL treatment, the fraction of second-best votes for the status quo is generally consistent with the mathematical predictions based on the value of $\pi$. Thus, strategic votes for the status quo as the second-best option happen a predictable percentage of the time.

*4.3. Result 3: Most choices of the status quo option are not examples of status quo bias, but rather, are demand-revealing choices.*

Table 7 reports the overall rates of demand revelation by choice experiment design methodology and $\pi$. Although there are some differences in rates of demand revelation by choice experiment design[6], the main message of the table is about how demand revelation differs by value of $\pi$. Overall, subjects who have low values of $\pi$ are significantly less likely to cast a demand revealing vote than other subjects ($p = 0.03$ for $\pi = 0.125$ vs. 0.50; $p = 0.02$ for $\pi = 0.125$ vs. 0.75). These subjects prefer the status quo to most of the other available choice options, and have many bad options in the choice sets that they face. These are precisely the subjects whom voting theory predicts are likely to vote strategically, which is what we observe in Table 7.

[6] The overall rate of demand revelation is significantly less in the UBAL treatment than the OOD treatment (p = 0.02). There are no other significant differences in rates of demand revelation across methods to create fractional factorial choice experiment designs.

| | Mean | Robust Std. Err. | 95% Conf. Int. |
|---|---|---|---|
| By Method to Create Fractional Factorial Choice Experiments | | | |
| OOD | 0.93 | 0.015 | [0.90, 0.95] |
| UBAL | 0.87 | 0.019 | [0.83, 0.91] |
| RAND | 0.90 | 0.019 | [0.86, 0.93] |
| By Fraction of Choice Options Preferred to the Status Quo ($\pi$) | | | |
| $\pi = 0.125$ | 0.85 | 0.024 | [0.81, 0.90] |
| $\pi = 0.50$ | 0.92 | 0.015 | [0.89, 0.95] |
| $\pi = 0.75$ | 0.92 | 0.016 | [0.89, 0.95] |

Table 7. Rates of Demand Revelation

| | | | |
|---|---|---|---|
| By Method to Create Fractional Factorial Choice Experiments | | | |
| OOD | 0.89 | 0.037 | [0.81, 0.96] |
| UBAL | 0.95 | 0.026 | [0.90, 1.00] |
| RAND | 0.93 | 0.027 | [0.87, 0.98] |
| By Fraction of Choice Options Preferred to the Status Quo ($\pi$) | | | |
| $\pi = 0.125$ | 0.99 | 0.003 | [0.99, 1.00] |
| $\pi = 0.50$ | 0.84 | 0.048 | [0.75, 0.94] |
| $\pi = 0.75$ | 0.56 | 0.079 | [0.41, 0.72] |

Table 8. Fraction of Status Quo Votes that are Demand Revealing

Although the results demonstrate that strategic votes for the second-best option fall on the status quo option in a predictable fraction of choices, it is also the case that most votes for the status quo option are not strategic, but rather, are demand-revealing. Table 8 reports the fraction of status quo votes that are demand revealing. This fraction differs significantly ($p < 0.01$ for all differences) by a subject's value of $\pi$ in exactly the direction predicted by the mathematics of combinatorial choice set design. Those subjects who have very few choice options that are better than the status quo (those with $\pi = 0.125$) frequently cast demand-revealing votes for the status quo, while those who have many choice options that are better than the status quo ($\pi = 0.75$) cast demand-revealing votes for the status quo much less often. Out of 1031 total votes for the status quo, 954 are demand-revealing. Of the 77 votes that remain, 54 (70%) are strategic votes for the status quo when it is the second-best option in the choice set. Thus, only 23 votes for the status quo, out of a total of 1031 total status quo votes cast, appear to exhibit what we might term "status quo bias." Thus, when the incentives of the choice mechanism are accounted for, status quo bias appears to be a rare phenomenon in choice experiments for a public good.

**5. Conclusion**

Moving from the laboratory to the field, the results reported above have clear implications for any application of choice experiment data in public decision-making. When choice experiment data are analyzed to generate estimates of consumer preferences for use in determining the appropriate allocation of public resources to competing projects or programs, normally, the data analysis hinges on two assumptions: (1) choices of the non-status quo option are demand-revealing; and (2) choices of the status quo-option are not true reflections of underlying preferences, but rather, reflect an underlying behavioral bias in favor of the status quo. These experimental results, which are robust to the types of choice experiment design methodologies tested, demonstrate that both assumptions are likely to be invalid. The only reason that non-demand revealing choices fall more often on the status quo option is because the status quo option appears in every choice set, and due to the mathematics of combinatorial choice set design, is likely to land in the first-best or second-best position in a choice set a large percentage of the time. Because it is easier to distinguish choices of the status quo from other choices in the data, the apparent over-choice of the status quo has been interpreted to reflect behavioral bias on the part of consumers. This experiment demonstrates that it is far more likely that consumers are making either a demand-revealing choice of their favorite option in the choice set, or a strategic choice of the second-best option in order to prevent a very bad third option from winning. Failure to account for the incentive structures inherent in the multiple choice voting mechanism for public goods is likely to result in mis-estimation of preferences, and hence, misallocation of public resources. Thus, we caution against the use of the choice experiment methodology as currently practiced to inform public policy decisions.

Based on these results, we recommend that, in the future, researchers focus their attention on developing new mechanism designs that are free of the incentive properties that we describe above, or alternative methods to analyze choice data that exploits the information that is present in a respondent's choice of the second-best option.

## References


Adamowicz, W., P. Boxall, M. Williams, J. Louviere (1998), Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation, *American Journal of Agricultural Economics 80*(1), 64-75.

Adamowicz. W., D. Dupont, A. Krupnick, J. Zhang (2011), Valuation of cancer and microbial disease risk reductions in municipal drinking water: An analysis of risk context using multiple valuation methods, *Journal of Environmental Economics and Management 61*(2), 213-226.

Aravena, C., P. Martinsson, R. Scarpa (2014), Does money talk?—The effect of a monetary attribute on the marginal values in a choice experiment. *Energy Economics 44*, 483-491.

Australian Energy Market Operator (2014), Value of customer reliability review: Final report. http://www.aemo.com.au/-/media/Files/PDF/VCR-final-report--PDF-update-27-Nov-14.pdf

Bateman, I.J., A. Munro, G.L. Poe (2008), Decoy effects in choice experiments and contingent valuation: Asymmetric dominance, *Land Economics 84*(1),115-127.

Brownstone, D., K. Train (1999), Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics 89*, 109-129.

Cameron, T. A., J. R. DeShazo (2013), Demand for health risk reductions, *Journal of Environmental Economics and Management* 65(1) 87-109.

Carlsson, F., P. Frykblom, C.J. Lagerkvist (2007), Preferences with and without prices – does the price attribute affect behavior in stated preference surveys? *Environmental and Resource Economics 38*,155-164.

Collins, J.P., C.A. Vossler (2009), Incentive compatibility tests of choice experiment value elicitation questions, *Journal of Environmental Economics and Management 58*(2), 226-235.

Day, B., I.J. Bateman, R.T. Carson, D. Dupont, J. Louviere, S. Morimoto, R. Scarpa, P. Wang (2012), Ordering effects and choice set awareness in repeat-response stated preference studies, *Journal of Environmental Economics and Management 63*, 73-91.

Day, B. , J-L Pinto Prades (2010), Ordering anomalies in choice experiments, *Journal of Environmental Economics and Management 59*(3), 271-285.

Emmerson, C., P. Metcalfe (2013), Southern water customer engagement (economic) – willingness to pay, Report prepared for Southern Water UK, https://www.southernwater.co.uk/Media/Default/PDFs/A05_WillingnessToPay.pdf (accessed 14 Feb 2019).

Farquharson, R. (1969), *Theory of Voting*. New Haven: Yale University Press.

Felsenthal D.S., A. Rapoport, Z. Maoz (1988), Tacit cooperation in three alternative non-cooperative voting games: A new model of sophisticated behavior under plurality procedure, *Election Studies 7*, 143-161.

Ferrini, S., R. Scarpa (2007), Designs with a priori information for nonmarket valuation valuation with choice experiments: A Monte Carlo study, *Journal of Environmental Economics and Management 53*, 342-363.

Forsythe, R., R. Myerson, T. Rietz, R.J. Weber (1993), An experiment on coordination in multi-candidate elections: The importance of polls and election histories, *Social Choice and Welfare 10*(3), 223-247.

Fujiwara, T. (2011), A regression discontinuity test of strategic voting and Duverger's Law, *Quarterly Journal of Political Science 6*, 197-233.

Hensher, D. A., J.M. Rose, W.H. Greene (2015). *Applied Choice Analysis, Second Edi*. ed.

Huber, J., K. Zwerina (1996), The importance of utility balance in efficient choice designs, *Journal of Marketing Research 33*(3), 307-317.

Kanninen, B.J. (2002), Optimal design for multinomial choice experiments, *Journal of Marketing Research 39*(2), 214-227.

Kawai, K., Y. Watanabe (2013), Inferring strategic voting, *American Economic Review 103*(2), 624-662. http://dx.doi.org/10.1257/aer.103.2.624

List, J.A., P. Sinha, M.H. Taylor (2006), Using choice experiments to value non-market goods and services: Evidence from field experiments, *Advances in Economic Analysis and Policy 5*(2), 1132. https://doi.org/10.2202/1538-0637.1132

Louviere, J.J. (1984), Using Discrete Choice Experiments and Multinomial Logit Choice Models to Forecast Trial in a Competitive Retail Environment: A Fast Food Restaurant Illustration, *Journal of Retailing 60*(4), 81-108.

Louviere, J.J. (1988), Conjoint Analysis Modeling of Stated Preferences: A Review of Theory, Methods, Recent Developments and External Validity, *Journal of Transport Economics and Policy 10*, 93-119.

Louviere, J.J., G. Woodworth (1983), Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data, *Journal of Marketing Research 20*, 350-367.

Lusk, J.L., T.C. Schroeder (2004), Are choice experiments incentive compatible? A test with quality-differentiated beef steaks. *American Journal of Agricultural Economics 86*(2), 467-482.

McFadden, D. (1986), The choice theory approach to market research. *Marketing Science 5*(4), 275-297.

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

Queens University Belfast and Perceptive Insight (2015), Discrete choice experiments for valuing the benefits of improved NIE services. Report prepared for Northern Ireland Electricity Networks. https://www.nienetworks.co.uk/documents/consultations/nie-wtp-report-5nov2015.

Revelt, D., K. Train (1998), Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics 80*(4), 647-657.

Ryan, M., S. Wordsworth (2000), Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. *Scottish Journal of Political Economy 47*(5), 504-524.

Street, D.J., L. Burgess, J.J. Louviere (2005), Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments, *International Journal of Research in Marketing 22*, 459-470. doi:10.1016/j.ijresmar.2005.09.003

Smith, J.S., M. McKee (2007), 'People or prairie chickens' revisited: Stated preferences with explicit non-market trade-offs, *Defence and Peace Economics 18*(3), 223-244.

Taylor, L.O., M.D. Morrison, K.J. Boyle (2010), Exchange rules and incentive compatibility of choice experiments, *Environmental and Resource Economics 47*(2), 197-220.