Electoral Studies 38 (2015) 28–37

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Using elicited choice probabilities in hypothetical elections to study decisions to vote

Adeline Delavande [a, b], Charles F. Manski [c, *]

[a] Institute for Social and Economic Research, University of Essex, UK
[b] Nova School of Business and Economics, Portugal
[c] Department of Economics and Institute for Policy Research, Northwestern University, USA

ARTICLE INFO

Article history:
Received 22 January 2014
Received in revised form 13 January 2015
Accepted 13 January 2015
Available online 22 January 2015

Keywords:
Probabilistic polling
Hypothetical choice
Voting decisions

ABSTRACT

This paper demonstrates the feasibility and usefulness of survey research asking respondents to report voting probabilities in hypothetical election scenarios. Posing scenarios enriches the data available for studies of voting decisions, as a researcher can pose many more and varied scenarios than the elections that persons actually face. Multiple scenarios were presented to over 4000 participants in the American Life Panel (ALP). Each described a hypothetical presidential election, giving characteristics measuring candidate preference, closeness of the election, and the time cost of voting. Persons were asked the probability that they would vote in this election and were willing and able to respond. We analyzed the data through direct study of the variation of voting probabilities with election characteristics and through estimation of a random utility model of voting. Voting time and election closeness were notable determinants of decisions to vote, but not candidate preference. Most findings were corroborated through estimation of a model fit to ALP data on respondents' actual voting behavior in the 2012 election.

1. Introduction

Social scientists have long struggled to understand why persons vote in large elections and why turnout varies across elections. See Aldrich (1993), Feddersen (2004), Geys (2006a, b) and Smets and van Ham (2013) for review articles.

When performing empirical research on voting, it is natural to think first of analyzing data on actual elections. Aggregate data on turnout at the district or other geographic level are readily available and occasionally enable creative analysis as natural experiments (e.g. Brady and McNulty, 2011). However, these data do not describe individual voters and hence are ordinarily not well-suited to study interpersonal variation in decisions to vote.

Surveys of individuals can provide richer data by asking persons to report their voting behavior, socioeconomic-demographic attributes, and their perceptions of election characteristics. However, surveys of voting in actual elections have significant limitations. First, persons typically face actual elections only once every two or four years. Second, there may not be much temporal variation in the characteristics of candidates and other aspects of actual elections. Third, although theories of voting commonly consider an idealized setting in which a person chooses whether to participate in an isolated election for a single office, actual decisions to vote usually occur in a complex environment with contemporaneous elections for multiple offices and possibly ballot initiatives as well.

Given these limitations of data on actual elections, we think it useful to also perform empirical studies that pose hypothetical election scenarios and ask persons how they would vote in these scenarios. Data of this type can overcome the three limitations of data on actual elections. The researcher can pose many more scenarios than the number of elections that persons actually face. The researcher can design

* Corresponding author.
E-mail address: cfmanski@northwestern.edu (C.F. Manski).

http://dx.doi.org/10.1016/j.electstud.2015.01.006
0261-3794/

the scenarios to exhibit considerable variation in the characteristics of candidates and other aspects of the election. And one can pose scenarios that hypothesize an isolated election.

Of course studies of voting in hypothetical elections are not a panacea. One concern is that the responses that persons give may differ from the way that they would actually behave. Another is that the scenarios that a researcher can pose in practice inevitably omit some features of the environment that a person would face in an actual election. These concerns are legitimate, but studies of hypothetical elections can still usefully add to the empirical evidence currently available for analysis of decisions to vote.

The broad precedent for our study is a long history of applied econometric research that poses choice scenarios, asks persons to state the choices they would make in these scenarios, and uses the data to estimate random-utility models of choice behavior, in the same manner that data on actual choices would be used. See, for example, Beggs et al. (1981), Fischer and Nagin (1981), Louviere and Woodworth (1983), Manski and Salomon (1987), and Ben-Akiva and Morikawa (1990).

Our specific precedents are the methodological and empirical studies of Manski (1999) and Blass et al. (2010). Manski (1999) reasoned that stated choices may differ from actual ones because researchers provide respondents with different information than they have when facing actual choice problems. The norm has been to pose *incomplete scenarios*, ones in which respondents are given only a subset of the information they would have in actual choice settings. When scenarios are incomplete, stated choices cannot be more than point predictions of actual choices.

Elicitation of *choice probabilities* overcomes the inadequacy of stated-choice analysis by permitting respondents to express uncertainty about their behavior in incomplete scenarios. Manski (1999) showed how elicited choice probabilities may be used to estimate random utility models with random coefficients. Blass et al. (2010) used the methodology to estimate consumer preferences for electricity reliability. The present paper uses it to estimate a random utility model of voting decisions, the data being voting probabilities in hypothetical elections.

The broad idea of measuring choice intentions probabilistically has much precedent, dating back to Juster (1966). See Manski (2004), Hurd (2009), and Delavande (2014) for review articles. Eliciting choice intentions probabilistically might be viewed as more cognitively demanding than eliciting them verbally, yet previous research has amply illustrated its feasibility by showing that most respondents are able to respond meaningfully in probabilistic terms when asked about events germane to their lives. Probabilistic measurement of voting intentions in actual elections has recently been implemented on a large scale in the American Life Panel (ALP). Delavande and Manski (2010, 2012) study the voting probabilities that ALP respondents reported prior to the 2008 presidential election and the 2010 congressional and gubernatorial elections. Kapteyn et al. (2014) study voting probabilities reported prior to the 2012 presidential election.

The present study differs from the above research using ALP data in two important respects. First, it analyzes data on voting probabilities in hypothetical elections rather than voting probabilities prior to actual elections. Second, it uses the data to study how the decision to vote varies with the characteristics of the election.

Section 2 describes the ALP, the design of the election scenarios, and the sample whose responses we analyze. Section 3 uses the data to examine the decision to vote. We first present suggestive findings on the univariate variation of voting probabilities with election characteristics. We next explain the structure and estimation of the random utility model. We then pose a particular model specification and present the parameter estimates. To close the empirical analysis, we compare the estimates with those of a similar model estimated using ALP data on respondents' actual voting decisions. Section 4 discusses what we have learned substantively about voting and methodologically about survey research posing hypothetical election scenarios.

2. Data description

2.1. The American Life Panel

The American Life Panel is a national longitudinal survey of Americans of age 18 and older, begun by RAND in 2003. Since its start, the ALP has expanded from about 500 to roughly 4500 respondents. The ALP recruits participants from several sources, including representative samples of the population and convenience samples.[1]

The ALP sampling process yields a wide spectrum of participants. However, respondents over-represent some demographic groups relative to others. The first column of Table 1 describes the composition of the 4329 participants who responded to at least one of the three survey waves that posed hypothetical election scenarios. These waves were conducted several weeks apart in November and December 2012 following the presidential election. Only U.S. citizens were invited to respond to the election questions.

Relative to the population of the United States, the participants were more often female (60 percent) and college educated (37 percent with 16 or more years of schooling compared to 28 percent in the 2010 census; U.S. Census Bureau, 2013). They were similar in terms of ethnic group (12 percent Black and 17 percent Latinos) and proportion of adults above age 65 (18 percent). Among panel members who participated in at least one of the three waves with hypothetical election scenarios, 60 percent participated in all waves and 24 percent in two.

2.2. The hypothetical election scenarios

ALP panel members participating in the three survey waves were asked their intention to vote in a set of hypothetical presidential elections. Each scenario presented several election characteristics: (i) how much the participant likes each of the candidates, as measured on a thermometer scale previously used in the American National Election Survey, (ii) the closeness of the election as measured by a poll, and (iii) how costly it may be to vote in term of time. As

[1] For details see https://mmicdata.rand.org/alp/index.php?page=panelcomposition.