INTERNATIONAL ECONOMIC REVIEW
Vol. 51, No. 2, May 2010

USING ELICITED CHOICE PROBABILITIES TO ESTIMATE RANDOM UTILITY MODELS: PREFERENCES FOR ELECTRICITY RELIABILITY*

BY ASHER A. BLASS, SAUL LACH, AND CHARLES F. MANSKI[1]

ERCG Ltd., Israel; The Hebrew University of Jerusalem, Israel, and CEPR, U.S.A.; Northwestern University, U.S.A.

When choice data are not available, researchers studying preferences sometimes ask respondents to state the actions they would choose in choice scenarios. Data on stated choices are then used to estimate random utility models, as if they are data on actual choices. Stated and actual choices may differ because researchers typically provide respondents less information than they would have in actuality. Elicitation of choice probabilities overcomes this problem by permitting respondents to express uncertainty about behavior. This article shows how to use elicited choice probabilities to estimate random utility models and reports estimates of preferences for electricity reliability.

1. INTRODUCTION

When suitable data on actual choices are not available, researchers studying consumer preferences sometimes pose hypothetical choice scenarios and ask respondents to state the actions they would choose if they were to face these scenarios. The data on *stated choices* are then used to estimate random utility models, in the same manner as are data on actual choices. 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).

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) sketched how elicited choice probabilities may be used to estimate random utility models with random coefficients. This article further develops the approach and reports the first empirical implementation.

The article is organized as follows. In Section 2, the random utility model usually used in the analysis of stated-choice data is presented. Assuming an extreme value distribution of the random utility term leads to the standard multinomial logit model and to the mixed-logit model when the utility function has random coefficients. In Section 3, the advantages of eliciting choice probabilities as opposed to stated choices are discussed, and it is shown that estimation of the utility function parameters is simpler and relies on weaker distributional assumptions.

* Manuscript received April 18, 2008; revised October 10, 2008.

[1] The data analyzed in this research were collected under a contract from the Israel Ministry of National Infrastructures to ERCG, Ltd. The research of Charles F. Manski was supported in part by U.S. National Science Foundation grant SES-0549544. We thank Yaron Aronshtam for his excellent research assistance. We thank three anonymous referees for many constructive comments. Please address correspondence to: Charles F. Manski, Department of Economics and Institute for Policy Research, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A. Phone: 847-491-8223. Fax: 847-491-7001. E-mail: *cfmanski@northwestern.edu*.

In Section 4, we apply the "elicited choice probability" methodology to estimate preferences for reliability in the supply of electricity to households in Israel. We describe the process of eliciting choice probabilities in hypothetical scenarios from a sample of households and analyze their responses. We then use the elicited choice probabilities to estimate mean preferences and willingness to pay (WTP) for reductions in the duration or frequency of electricity outages. We also present estimates of individual preferences that are derived from the model in a straightforward manner. Conclusions close the article.

2. ECONOMETRIC ANALYSIS OF STATED CHOICES

Let i denote an individual asked to respond to a choice scenario. In standard stated-choice analysis, the respondent is presented with J hypothetical alternatives, $j = 1, \ldots, J$, and is asked to choose one. Let y_i denote the stated choice. Let person i have observed attributes s_i. Let each alternative j presented to person i have stated characteristics $v_{ij.}$ For example, in a study of preferences for electricity reliability, the alternatives may differ in the stated duration and frequency of electricity outages and in the price of electricity.

It is common to assume that the utility of alternative j has the random-coefficients form

$$U_{ij} = x_{ij}\beta_i + \varepsilon_{ij}. \tag{1}$$

Here $x_{ij} = x(v_{ij}, s_i)$ is a specified function of observed alternative characteristics and personal attributes, and ε_{ij} is a utility component that is observed by the decision maker but not by the researcher. Let $x_i \equiv (x_{ij,}\ j = 1, \ldots, J)$. It is also common to assume that $\varepsilon_i \equiv (\varepsilon_{ij}, j = 1, \ldots, J)$ are independent and identically distributed (i.i.d.) conditional on x_i, with the Type I extreme value distribution. Then the probability of stating choice j conditional on x_i and β_i has the multinomial logit form

$$P(y = j \mid x_i, \beta_i) = \frac{e^{x_{ij}\beta_i}}{\sum_{h=1}^{J} e^{x_{ih}\beta_i}}. \tag{2}$$

Finally, assume that β is statistically independent of x, with density $f(\beta \mid \theta)$ across the population of potential respondents, where the form of f is known up to the parameter vector θ. This yields the mixed-logit model of McFadden and Train (2000)

$$P(y = j \mid x_i, \theta) = \int \frac{e^{x_{ij}\beta}}{\sum_{h=1}^{J} e^{x_{ih}\beta}} f(\beta \mid \theta) d\beta. \tag{3}$$

These choice probabilities provide the basis for estimation of the parameters θ by maximum likelihood or another method.

3. ECONOMETRIC ANALYSIS OF ELICITED CHOICE PROBABILITIES

3.1. *Elicited Choice Probabilities and Stated Choices.* Stated-choice analysis assumes that individual i responding to a choice scenario knows the value of both x_i and ε_i and, hence, is able to state a definite utility-maximizing choice. However, the researcher only provides the stated characteristics $v_i \equiv (v_{ij}, j = 1, \ldots, J)$, which determine x_i but not ε_i. It is therefore questionable that the respondent knows ε_i when the scenario is posed.

Eliciting choice probabilities enables respondents to express uncertainty about ε_i. It permits a person to treat ε_i as a vector of utility components whose value need not be known when responding to the choice scenario, but which would be known in an actual choice setting. Formally, suppose that person i forms a subjective distribution for ε_i, derives the subjective probability that he would choose each alternative in an actual choice setting, and reports these subjective probabilities to the researcher. Let q_{ij} denote the choice probability reported by person i for alternative j. Then q_{ij} is the subjective probability that person i places on the event that the realizations of ε_i will make option j optimal.

Suppose, in particular, that person i has utility function (1) and, given the stated characteristics v_i, places a continuous subjective distribution Q_i on ε_i. Then his subjective choice probability for alternative j is

$$q_{ij} = Q_i[x_{ij}\beta_i + \varepsilon_{ij} > x_{ik}\beta_i + \varepsilon_{ik}, \quad \text{all } k \neq j]. \tag{4}$$

The right-hand side of Equation (4) gives a subjective random utility interpretation of elicited choice probabilities.

Subjective distribution Q_i expresses *resolvable uncertainty*, that is, uncertainty about utility components that are not stated in the choice scenario but that would be known in an actual choice setting. A person contemplating a choice scenario may also face *unresolvable uncertainty*. That is, there may be utility components that the person believes would remain unknown in an actual choice setting. The usual economic assumption is that a person copes with unresolvable uncertainty by maximizing subjective expected utility. If a person only faces unresolvable uncertainty when responding to the choice scenario, he would place subjective probability one on the alternative that maximizes expected utility.

It is important to understand how elicited choice probabilities and stated choices are related to one another. Assuming that persons form subjective distributions in the manner described earlier, Juster (1966) and Manski (1990) reasoned that a person asked a stated-choice question computes his subjective choice probability for each alternative and reports the one with the highest probability. Thus, when person i states that he would choose alternative j, he means that $q_{ij} \geq q_{ik}$, all $k \neq j$. He does not necessarily mean that $U_{ij} \geq U_{ik}$, all $k \neq j$, as assumed in standard stated-choice analysis. The latter assumption is essentially correct only when $q_{ij} = 1$, which means that person i places subjective probability one on the event $(U_{ij} \geq U_{ik}, \text{all } k \neq j)$.[2]

The above mentioned reasoning implies that the standard derivation of the mixed-logit model, described in Section 2, is valid only when the subjective choice probabilities of all respondents take the extreme values zero and one. Suppose to the contrary that a group of respondents places a nonextreme subjective probability, say 0.6, on choosing alternative j. If asked to state their choices, all of these respondents would state that they would choose j, and the standard derivation assumes that all of them would choose j in actuality. However, if respondents have accurate expectations and if their realizations of ε are statistically independent, then only 0.6 of them would actually choose j.

Eliciting choice probabilities is more informative than asking for stated choices. When a person is sure that he would choose option j, he can express this belief by placing probability one on j. When he is uncertain whether he would choose j, he can report a nonextreme choice probability.

A person reporting nonextreme probabilities expresses the belief that, in an actual choice setting, he would possess choice-relevant information beyond the characteristics v_i provided to him in the choice scenario; that is, some of the uncertainty he faces is resolvable. Equation (4) quantifies the information that person i anticipates learning in an actual choice setting. If he believes that he would acquire no further information in an actual setting, then Q_i is degenerate and he reports extreme choice probabilities. If he believes that he might obtain information that would affect his choice, then Q_i is nondegenerate and he reports nonextreme choice probabilities.

[2] We say that the assumption is "essentially correct" because the utility inequalities hold with subjective probability one, not with certainty.

3.2. *The Linear Mixed-Logit Subjective Random Utility Model.* In order to use elicited choice probabilities to estimate the subjective random utility model (4) requires assumptions on the subjective distribution Q_i that each respondent i places on ε_i and assumptions on the cross-sectional distribution of the random coefficients β. This subsection and the next pose assumptions similar to those maintained in standard stated-choice analysis. The result is an extremely simple form of econometric analysis. Subsection 3.4 considers weaker assumptions that may be more credible.

The standard practice in stated-choice analysis has been to assume that the components of (ε_{ij}, $j = 1, \ldots, J$) are objectively i.i.d. with the extreme value distribution. Suppose that respondents make the same assumption subjectively. Then the choice probabilities (4) have the multinomial logit form

$$q_{ij} = \frac{e^{x_{ij}\beta_i}}{\sum_{h=1}^{J} e^{x_{ih}\beta_i}}, \quad j = 1, \ldots, J. \tag{5}$$

Applying the log-odds transformation to Equation (5) yields the linear mixed-logit model

$$\ln\left(\frac{q_{ij}}{q_{i1}}\right) = (x_{ij} - x_{i1})\beta_i = (x_{ij} - x_{i1})b + u_{ij}, \quad j = 2, \ldots, J \tag{6}$$

where $\beta_i = b + \eta_i$, $u_{ij} = (x_{ij} - x_{i1})\eta_i$, and the alternative designated $j = 1$ is arbitrarily chosen.

Standard analysis of stated choices assumes that the cross-sectional distribution of β, hence η, is statistically independent of x. Let this assumption hold.[3] Without loss of generality, set $E(\eta) = 0$ as a normalization. It then follows that $b = E(\beta)$, $E(u|x) = 0$, and Equation (6) is the linear mean regression model

$$E\left[\ln\left(\frac{q_{ij}}{q_{i1}}\right) \middle| x\right] = (x_{ij} - x_{i1})b. \tag{7}$$

3.3. *Rounding and Symmetry of Preferences.* If model (7) is taken literally, the mean-preference parameters b may be consistently estimated by least squares, without need to assume anything about the shape of the distribution of β. This contrasts with standard econometric analysis of stated choices, where the researcher must specify a parametric family of distributions for β.

However, we cannot take the model quite literally. As will be discussed in Section 4, respondents tend to round their responses to the nearest 5–10%. Such minor rounding has been found to be commonplace in elicitation of subjective probabilities; see Manski (2004) and Manski and Molinari (2010).

Rounding of interior subjective probabilities (say, from 43% to 45%) is relatively unproblematic. However, rounding of values near zero and one raises a serious difficulty due to the sensitivity of the log odd function near the boundaries of the [0, 1] interval. At the extreme, some respondents report subjective choice probabilities equal to zero or one, thus generating log odds that equal minus or plus infinity. Hence, least squares estimation breaks down.[4]

The inference problem created by rounding of small (large) subjective probabilities to zero (one) can be resolved if preferences are symmetrically distributed with center at b. Symmetry of

[3] In the application of Section 4, the attributes v of the hypothetical alternatives are randomly drawn; hence, η is statistically independent of v. We will assume that η is statistically independent of the observed personal attributes s that are determinants of x; hence, η is statistically independent of x in our application.

[4] One should not drop the cases with choice probabilities equal to zero or one, because this truncates the sample in a response-based manner. One might consider an ad hoc transformation of reported zeroes and ones to values near these boundaries, but the least squares estimates are sensitive to the transformation performed.

preferences has been a common assumption in stated-choice analysis, which typically supposes that β has a normal distribution. Symmetry implies that the unobserved u_{ij} are symmetrically distributed about zero conditional on x_i and, hence, have median zero conditional on x_i. Thus, we have the linear median regression model

$$M\left[\ln\left(\frac{q_{ij}}{q_{i1}}\right)\middle| x\right] = (x_{ij} - x_{i1})b, \tag{8}$$

whose parameters may be estimated by least absolute deviations (LAD) in the absence of rounding.

A well-known robustness property of the median of a random variable is its invariance to transformations that do not alter the ordering of values relative to the median. Thus, if y is a random variable with median M, then M is also the median of any function $f(y)$ such that $y < M \Rightarrow f(y) < M$ and $y > M \Rightarrow f(y) > M$. This holds even if the function f transforms small values of y to $-\infty$ and large ones to ∞. Hence, Equation (8) continues to be the same linear median regression if small values of q are replaced by zero and large values by one.

In the application of Section 4, we will assume that preferences are symmetrically distributed. When a reported subjective probability is zero or one, we transform it to a value close to zero or one. We then estimate model (8) by LAD, whose result is insensitive to the specific transformation used. For brevity, we will continue to refer to b as the "mean preferences" rather than the more cumbersome "center of symmetry of the preference distribution."

3.4. *Models with Weaker Assumptions on Respondent Beliefs.* The linear mixed-logit model derived earlier has many appealing properties. The basic form of the model presented in Subsection 3.2 places no restrictions on the population distribution of preference parameters and is easily estimable by least squares. The robust-to-rounding version of the model developed in Subsection 3.3 requires only that the distribution of preferences be symmetric and be easily estimable by LAD.

These properties compare favorably with standard econometric analysis of stated choices. As discussed in Subsection 3.1, the standard approach is valid only when respondents are certain of their stated choices. It requires specification of a parametric family for the distribution of preferences. Moreover, it requires often difficult numerical maximization of a likelihood function or solution of a set of nonlinear moment equations.

Perhaps the main unappealing feature of the linear mixed-logit model is its assumption that each respondent i believes $(\varepsilon_{ij}, j = 1, \ldots, J)$ to be i.i.d. with the extreme value distribution. This distributional assumption is highly convenient. As a consequence, it has been prominent as an assumption on the objective distribution of ε from the original conditional logit model of McFadden (1974) through the more recent mixed-logit model of McFadden and Train (2000). However, convenience does not make an assumption credible. Previous applications of the assumption have not been able to motivate it persuasively. Neither can we.

Cognizant of the problem, econometricians have over the years developed numerous approaches to discrete choice analysis that rest on more credible, albeit less convenient, assumptions about objective probability distributions; see, for example, Manski (1975, 2007), Horowitz (1992), and Matzkin (1992). These approaches can also be applied in the present setting, which requires assumptions on subjective distributions. We discuss one here.

3.4.1. *Maximum score estimation of subjective random utility models*. Considering a binary choice setting with alternatives (j, k), Manski (1999) suggested assuming only that each person i places subjective median zero on $\varepsilon_{ij} - \varepsilon_{ik}$ and that the cross-sectional distribution of β is symmetric. The first assumption yields the inequality

$$q_{ij} \geq 0.5 \Leftrightarrow (x_{ij} - x_{ik})\beta_i \geq 0. \tag{9}$$

The second assumption yields the inequality

$$P(q_{ij} \geq 0.5 \mid x) \geq 0.5 \Leftrightarrow (x_{ij} - x_{ik})b \geq 0. \tag{10}$$

Inequality (10) can be exploited to estimate b by the maximum score method (Manski, 1975, 1985).

Indeed, inequality (10) may be applied to stated-choice data as well as to elicited choice probabilities. As discussed in Subsection 3.1, a person's statement that he would choose option j over k means that $q_{ij} \geq 0.5$. Hence, the maximum score method can be used to estimate a subjective random utility model with stated-choice data. In contrast with the standard approach, one need not assume that respondents are certain about their stated choices.

These nice features of maximum score estimation must be balanced against the fact that inequality (10) point-identifies b only if the attributes x have sufficiently rich support, as explained in Manski (1988) and elsewhere. Researchers typically present only a finite set of distinct choice scenarios to respondents, which do not suffice for point identification. One may, nevertheless, exploit (10) to bound b. The maximum score method consistently estimates the bounds. We give an illustrative application in Subsection 4.7.

4. ESTIMATING PREFERENCES FOR ELECTRICITY RELIABILITY

We have applied the "elicited-choice-probabilities" approach to estimate consumer valuation of residential electricity reliability in Israel. Knowledge of consumer WTP for reliability is an important component of a rational planning strategy for capacity investment in the generation and transportation of electricity, as well as a key factor in determining an optimal electricity pricing schedule.

Subsection 4.1 cites previous studies using the stated-choice approach. Subsection 4.2 describes our research design. Subsection 4.3 reports basic findings on the elicited choice probabilities. Subsection 4.4 presents estimates of the mean consumer preferences b in models of form (8). Subsection 4.5 derives estimates of willingness-to-pay for electricity reliability by persons with the mean preferences. Subsection 4.6 presents findings on the dispersion of preferences. Finally, Subsection 4.7 uses the maximum score method to estimate a simple version of the model with weakened assumptions discussed in Subsection 3.4.

4.1. *Stated-Choice Studies.* Various theoretical and applied models have been developed in the resource and energy literature to estimate the marginal value of service reliability.[5] Models have often been estimated using stated-choice data, with customers asked to choose among different bundles of service attributes. Important stated attributes include the number and duration of outages as well as the cost of service. Revelt and Train (1998) and Goett et al. (2000) used stated-choice data to estimate mixed-logit models of the type described in Equation (3). Cai et al. (1998) used a different methodology—an extension of the "double-bounded" procedure used in studies of contingent valuation of natural resources—to estimate WTP for electricity service attributes.

A previous study of the Israeli electricity market was conducted by Beenstock et al. (1998). They used both conjoint analysis and contingent-valuation data to estimate the willingness of Israeli households to pay for reduced power outages. They found the conjoint analysis more reliable. Here, they asked consumers to rank the hypothetical alternatives and they analyzed the rank-ordered data. Their estimates indicate that the perceived cost of unsupplied electricity to Israeli households in 1991 was between \$2.3 per KWh (in the spring/autumn during morning/midday hours) and \$11 per KWh (in the winter during afternoon/evening hours).

[5] See Caves et al. (1990) for a comparison of the different approaches and Lawton et al. (2003) for a recent review of U.S. results.

4.2. *Study Design.* In this study, a stratified random sample of 557 Israeli households was drawn and an adult member of each household was interviewed in person about his or her preferences for reliability of home electricity supply. The interviews were performed in summer 2005 by a professional survey research firm, with responses recorded on laptop computers using Computer Assisted Personal Interview (CAPI) software.

Each respondent was asked to report choice probabilities in ten different "games." A game specifies a scenario in which the person is presented with two alternative bundles of attributes ($J = 2$) and is asked to state his or her chances of choosing each alternative.

The games were introduced as follows:

"In each game we will show you two alternatives. Each alternative is characterized by a different bi-monthly electricity bill and by different numbers of outages and different average duration of the outages. In each game, you should evaluate what is the chance in percentage terms of choosing one of the two alternatives. The chance of each alternative should be a number between 0 and 100 and the chances given to the two alternatives should add up to 100. For example, if you give a 5% chance to one alternative it means that there is almost no possibility that you will choose that alternative. On the other hand, if you give an 80% or over chance to an alternative it means that almost surely you will choose it."

Since the early 1990s, economists have developed considerable experience similarly asking respondents to state expectations of future events as the "percent chance" that the event will occur; Manski (2004) reviews the literature. The novelty in this study is our use of responses to this type of question as a replacement for actual choice data to estimate the parameters of a random utility model.[6]

In each game, the alternatives presented to sample members differed in the duration (D) and frequency (F) of outages and in the corresponding electricity bill (C). Thus, a stated alternative is a (D, F, C) triple. The stated frequencies of outages were 0 (perfect reliability), 1, 2, 4, 5, and 8 outages per season. The duration of an outage was stated as an interval of length 0–10, 10–60, 60–120, and 120–240 minutes. In our analysis, we took the duration to be the mid-point of the interval, that is, 5, 35, 90, or 180 minutes. The stated electricity bill (C) took one of five values: the household's actual electricity bill, 15% or 40% above the actual bill, or 10% or 20% below the actual bill. The stated combinations of (D, F, C) were randomly generated, but dominated alternatives were excluded.

The value that a person attaches to electricity reliability may depend on the timing of outages. For example, an anticipated outage at 11 a.m. on a weekday in the summer may be valued differently than an unanticipated outage at 8 p.m. on a weekend day in the winter. In order to assess the variation of valuation with timing, each game was set in a scenario characterized by three timing variables. The first timing variable is season: 50% of the games were set in the summer, 41% in the winter, and 9% were set in the spring (fall is very short in Israel). The second variable relates to the time of day when outages occur: during peak hours (39%), off-peak hours (30%), or intermediate hours (31%). The third timing variable refers to the days of the week, specifically whether outages occur during the weekend or not. An additional characteristic of the scenario is the availability of advance warning on the outages. This feature was present in only 9% of the games.

In order to illustrate, here is how a typical game was described

"Suppose you are in the summer and all the outages occur between 5PM and 10PM (peak hour in summer) during weekends and there is no advance warning. In alternative 1 you are offered 5 outages lasting up to 10 minutes each and the electricity bill will be 360 NIS. In alternative 2 you are offered 0 outages and the electricity bill will be 480 NIS."

The abbreviation NIS stands for New Israeli Shekel, the currency unit. One NIS was worth about 22 U.S. cents when the survey was administered in the summer of 2005.

[6] Delavande (2008) uses a different type of probabilistic expectations data to estimate a random utility model. She has data on actual choices and uses elicited expectations to characterize the attributes of alternatives.

TABLE 1

DISTRIBUTION OF SCENARIOS

Scenario	Count	Percent
No advance warning, off-peak, weekday, winter	587	10.8
No advance warning, peak, weekday, summer	584	10.8
No advance warning, intermediate, weekend, winter	568	10.5
No advance warning, intermediate, weekday, summer	569	10.5
No advance warning, off-peak, weekend, winter	524	9.7
No advance warning, off-peak, weekday, summer	531	9.8
No advance warning, peak, weekday, winter	538	9.9
No advance warning, intermediate, weekend, summer	537	9.9
No advance warning, peak, weekday, spring	499	9.2
Advance warning, peak, weekday, summer	493	9.1
Total	5,430	100

The games were played in 10 out of the 36 possible scenarios. These were the scenarios deemed to be of greatest interest and/or relevance for analysis. (For example, all weekend hours were considered off-peak.) In order not to confuse the respondent by changing the timing scenario in each game, the same scenario was maintained during five successive games. Thus, a respondent plays five different games in each of two scenarios. The scenarios changed across respondents to cover the 10 timing possibilities of interest. These scenarios and their frequency in the sample are listed in Table 1.

We believe the scenarios posed to be realistic in the Israel context. There has been an ongoing public debate in Israel about problems in generating the amount of electricity needed to satisfy growing demand. During the summer, in particular, newspapers are full of alarming news about warnings from the electricity company that it will not be able to generate enough electricity. Outages occur with some regularity and have occurred even more frequently in the past. Thus, we think that the survey respondents are not surprised when presented with scenarios where outages occur in different frequencies and durations.

4.3. *The Elicited Choice Probabilities.* The response rate to the questions posed was very high. Of the 557 sample members, 500 (89.8%) reported choice probabilities in all 10 games and 38 (6.8%) gave responses in 9 games. In all, 5,430 responses were obtained from the 557 sample members. Thus, the overall item response rate was 5430/5570 = 0.975.

Let q_{i1} and q_{i2} denote the choice probabilities (percent chance divided by 100) for the two alternatives elicited from individual i in a given game. In all cases, we have $q_{i2} = 1 - q_{i1}$. That is, the elicited choice probabilities always add up to 1.

The elicited percent-chance responses for alternative 1 are tabulated in Table 2 and their histogram is shown in Figure 1. The table and figure show that most responses are multiples of 10 and almost all of the rest are multiples of 5. The bimodality of the histogram in the figure reflects the random assignment of the stated characteristics of alternatives.

Minor rounding of responses is inconsequential for our analysis, but gross rounding would be problematic. It is therefore important to observe that the data do not show evidence of gross rounding. In particular, the elicited choice probabilities are not concentrated at the values 0, 50, and 100; the frequencies of these responses are just 0.08, 0.05, and 0.08, respectively. Instead, the distribution of responses is weakly bimodal. The most prevalent responses were 20% and 80%, with response frequencies of 0.118 and 0.119.

Table 2 shows the importance of eliciting choice probabilities rather than stated choices. A stated-choice question only permits the respondent to state a 0% or 100% chance of choosing an alternative. When given the opportunity to state a value in the [0, 1] interior, sample members state an interior value in 0.84 of their responses. Thus, respondents find the stated scenario too

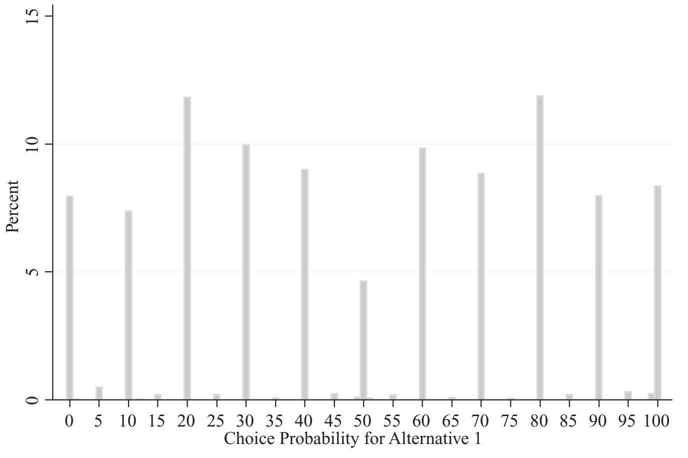

FIGURE 1

HISTOGRAM OF ELICITIED PROBABILITIES

incomplete to give a definitive choice response in at least 0.84 of the cases. We say "at least 0.84" because some of the observed responses of 0% and 100% may be rounded versions of interior values.

4.4. *Estimation of Mean Preferences.* Table 3 presents LAD estimates of the parameters b in model (8) for three specifications of the attributes x. In order to compute these estimates, we replace reported zero probabilities by 0.001 and ones by 0.999. As discussed in Section 3, LAD estimation of a median regression function is insensitive to the specific values that we use in place of zero and one.

The estimates in Table 3 use 3,947 of the available 5,430 observations, and comprise data from 556 of the 557 households. (The one deleted household participated in only a single game, which was dropped for one of the reasons described here.) The reduction in sample size occurred for two reasons. First, we had to drop 990 games where the software erroneously generated alternatives in which outages occur with positive frequency but zero duration. We were concerned that respondents might be confused by such alternatives, so we dropped the games in which they occurred.

Second, we dropped the 493 observations on games played in scenarios with advance warning of electricity outages. The reduction in consumer welfare caused by outages may depend on whether advance warning is given; hence, the utility parameters b should be interacted with this feature of the scenario. However, our sample of 493 observations with advance warning was too small to permit precise inference. We therefore decided to focus on the scenarios with unannounced outages, for which we have many more observations. The latter also are the scenarios of most interest to policy makers in Israel.

Recall that the unobserved component of utility in Equation (6) has the form $u_{ij} = (x_{ij} - x_{i1})\eta_i$. Thus, the random parameter specification implies that u is heteroskedastic and that it is correlated across the games played by a given sample member. These features of u do not affect the consistency of LAD estimation but do affect statistical inference.

We obtained the standard errors of the parameter estimates shown in Table 3 by cluster bootstrapping the sample. Cluster bootstrapping means that, to generate a pseudo-estimate of the parameters, we drew 556 respondents with replacement from the actual sample of 556 respondents and used the data on all of the games played by these persons to reestimate the

TABLE 2
TABULATION OF CHOICE PROBABILITIES

	(For Alternative 1)		
	Frequency	Percent	Cum.
0	432	8.0	8.0
1	1	0.0	8.0
5	26	0.5	8.5
10	400	7.4	15.8
12	1	0.0	15.8
15	11	0.2	16.0
20	642	11.8	27.9
25	11	0.2	28.1
30	541	10.0	38.0
35	3	0.1	38.1
40	488	9.0	47.1
45	12	0.2	47.3
49	5	0.1	47.4
50	252	4.6	52.0
51	3	0.1	52.1
55	10	0.2	52.3
60	534	9.8	62.1
65	4	0.1	62.2
70	480	8.8	71.0
75	2	0.0	71.1
80	645	11.9	82.9
85	11	0.2	83.1
90	433	8.0	91.1
95	17	0.3	91.4
99	12	0.2	91.6
100	454	8.4	100.0
Total	5,430	100	

model. We repeated this process 500 times to generate 500 pseudo-estimates of β. The reported standard errors are the standard deviations of these 500 pseudo-estimates.

4.4.1. *Parameter estimates.* The three columns of Table 3 present estimates of successively richer utility models. Column (1) reports estimates of a simple model that takes the attributes of an alternative to be D, F, and C. We further allow the electricity-cost coefficient to vary by income group. Respondents self-reported whether their household income is below the Israeli average income (36%), at the average income (21%), or above average (25%). The remaining 18% refused to answer the question and were assigned to a residual group. We would have preferred to ask respondents to report their household incomes, but this aspect of the questionnaire design was not under our control. Respondents were not told the average Israeli income. Hence, there is reason to question the accuracy of the self-reports of being below or above average.

In any case, the parameter estimates have the expected negative signs and the standard errors indicate that they are statistically precise. Observe that the marginal disutility of electricity cost is smaller for higher income households.

This first model makes the unrealistic assumptions that (a) the disutility of an additional outage is the same regardless of its duration and (b) the disutility of an additional minute per outage is the same regardless of the number of outages. In order to enable a more flexible description of utility, the model in column (2) adds to x the total outage time $T = F \times D$. This gives the following specification of the utility function:

$$U_{ij} = (\beta_{iC} \times IncomeGroup)\, C_j + \beta_{iD} D_j + \beta_{iF} F_j + \beta_{iT} T_j + \varepsilon_{ij}. \tag{11}$$

TABLE 3
LAD ESTIMATES OF UTILITY FUNCTION PARAMETERS*

Dependent Variable: Log Probability Ratio	(1)	(2)	(3)
Cost: baseline	−0.00142	−0.00163	−0.00148
	[0.00032]	[0.00030]	[0.00032]
Dummy for average income	0.00076	0.00088	0.00073
	[0.00035]	[0.00035]	[0.00036]
Dummy for above average income	0.00136	0.00151	0.00128
	[0.00035]	[0.00036]	[0.00038]
Dummy for missing income	0.00038	0.00055	0.00048
	[0.00045]	[0.00045]	[0.00046]
Duration	−0.00467	−0.00082	0.00122
	[0.00047]	[0.00108]	[0.00189]
Frequency	−0.1182	−0.0113	−0.0049
	[0.0149]	[0.0300]	[0.0565]
Total outage time ($T = F \times D$)	–	−0.00092	−0.00080
		[0.00020]	[0.00041]
Duration × weekend	–	–	−0.0039
	–	–	[0.0022]
Duration × peak			−0.0032
			[0.0024]
Frequency × weekend	–	–	−0.0395
			[0.0646]
Frequency × peak	–	–	0.0233
			[0.0676]
T × weekend	–	–	0.00009
			[0.0005]
T × peak	–	–	−0.00022
			[0.00049]
Observations	3,947	3,947	3,947

*Cluster-bootstrapped standard errors based on 500 replications are in brackets.

This specification extends Beenstock et al. (1998), who used only attribute T, without separate appearance of F and D.

Comparing the estimates in columns (1) and (2) shows that adding total outage time T to the model has essentially no effect on the electricity-cost coefficients, but it sharply reduces the magnitudes of the coefficients on D and F, which are now statistically indistinguishable from zero. The new coefficient on T is negative and is statistically precise. Nevertheless, we include D and F in the model since they allow for asymmetric marginal effects of D and F. Recall that we will use the estimated coefficients to compute the WTP for reductions in outages. The standard errors we report for WTP will account for the imprecision of the estimated coefficients on F and D.

The models in columns (1) and (2) presume that consumer valuation of electricity reliability does not depend on the timing of outages. As described in Section 3, the scenario specified for each game contains three timing indicators: season, hour, and weekend/weekday. In principle, the reduction in consumer welfare caused by outages may vary with these aspects of timing.

In order to explore this matter, we interacted D, F, and T with timing indicators. For simplicity, we grouped the off-peak and intermediate hours into a single off-peak category. We did not foresee any reason why the electricity-cost coefficients should vary with the timing of outages and, hence, did not interact C with the timing indicators.

F-tests of the interactions do not reject the hypothesis that the utility effect of outages does not vary across seasons (p-value 0.11). This result is in line with the general opinion of experts in the area and with the fact that there is little seasonal difference in household energy consumption in Israel.

With these preliminary results in mind, we present in column (3) a model that interacts D, F, and T with dummy variables that distinguish peak from off-peak weekday hours and weekdays relative to the weekend. Comparison of the estimates in columns (2) and (3) shows that addition of the interactions does not materially change the coefficients of the model in column (2), which now give the reduction in welfare associated with off-peak weekday outages. Notice that the estimated interactions for D and F are much larger than those for T (relative to their off-peak weekday coefficients). Individually, these interactions effects are not precisely estimated but, jointly, they cannot be rejected.[7] In any event, the standard error of the WTP will reflect the imprecision of the interaction effects.

4.4.2. *Parameter estimates for subpopulations.* It may be that households with different observed characteristics have systematically different preferences and expectations. In order to investigate this possibility, we split the sample into subsamples with different characteristics and estimated a separate model for each subsample. We performed three splits of the sample, these being by age, education level, and income, respectively. In each case, we estimated the model in column (3) of Table 3. Appendix Table A.1 presents the estimates.

We first split the sample between households where the respondent was 50 years of age or older and those where the respondent was less than 50 years old. There are 1,739 observations from 244 "old" households and 2,208 observations from 312 "young" households. Comparing the two groups, the estimates of the coefficients on F, D, T, and their interactions with the timing variables do not significantly differ from one another (p-value 0.45). The only statistically significant difference between the age groups is in the coefficients of the cost variable (C) and its interactions with the income group (p-value 0.027).

We next split the sample according to the respondent's education level. There are 179 households where the respondent finished college and 377 households where the respondent did not. Comparing the two groups, the estimates of the coefficients of F, D, T, and their interactions with the timing variables again do not significantly differ from one another (p-value 0.58). In this case, the cost coefficients do no differ between the education groups either (p-value 0.28).

Finally, we split the sample into three income groups: below average income (198 households), average income (119 households), above average (140 households), and households that did not provide this information (99 households). In this case, there are statistically significant differences in the coefficients of D and F and their interactions with the timing variables (p-value 0.004 and 0.08, respectively).

In sum, there is some evidence that preferences and expectations may vary with household characteristics such as age and income, but not education. For simplicity, in what follows we ignore these differences.

4.4.3. *Ordering effects.* Survey researchers and experimental economists are often concerned about the possibility of ordering effects in responses when persons are asked sequences of questions. Persons may tire as they proceed through the questions, but they also may become more familiar with the mode of questioning. Hence, the thoughtfulness of the responses may diminish or improve as the questioning proceeds.

Recall that each of our households faced five games in each scenario and that the scenarios were played one after the other. Ordering effects could occur within each scenario and across scenarios. In order to investigate these possibilities, we considered two definitions of "early" and "late" games. For each definition, we separately estimated the model of Table 3, column (3), using data from the early and late games. In one definition, we defined games 1–3 within a scenario to be the "early" games and games 4–5 to be the "late" games. In the other, we defined the five games of the first scenario to be the early games and those of the second scenario to be the late games.

[7] An F-test for zero interactions between "peak" and D, F, and T has p-value less than 0.01, whereas a similar test for the interaction with D and F only has p-value 0.03. Similarly, an F-test for the interactions between "weekend" and D, F, and T has p-value less than 0.01, whereas a similar test for the interaction with D and F only has p-value 0.12.

TABLE 4

MEAN WTP (US DOLLARS) FOR A ONE-MINUTE REDUCTION IN OUTAGE TIME* (HOUSEHOLDS WITH AVERAGE INCOME)

Reduction in D	(1) Weekday, Peak	(2) Weekday, Off-Peak	(3) Weekend, Off-Peak
$F = 1$	0.88	−0.12	0.98
	(0.49)	(0.45)	(0.48)
$F = 2$	0.59	0.05	0.60
	(0.25)	(0.18)	(0.24)
$F = 3$	0.49	0.11	0.47
	(0.17)	(0.10)	(0.17)
$F = 4$	0.45	0.14	0.40
	(0.14)	(0.08)	(0.13)
$F = 5$	0.42	0.16	0.36
	(0.13)	(0.07)	(0.11)
Reduction in F	Weekday, Peak	Weekday, Off-Peak	Weekend, Off-Peak
$D = 10$	−0.24	0.38	1.51
	(1.08)	(1.56)	(1.13)
$D = 30$	0.12	0.28	0.64
	(0.31)	(0.46)	(0.35)
$D = 60$	0.21	0.26	0.42
	(0.13)	(0.19)	(0.17)
$D = 90$	0.24	0.25	0.35
	(0.08)	(0.11)	(0.11)
$D = 120$	0.25	0.25	0.31
	(0.07)	(0.08)	(0.09)
$D = 150$	0.26	0.24	0.29
	(0.07)	(0.08)	(0.08)
$D = 180$	0.27	0.24	0.28
	(0.08)	(0.08)	(0.08)
$D = 300$	0.28	0.24	0.25
	(0.09)	(0.09)	(0.08)

*Cluster-bootstrapped standard errors are in parentheses.

We found that splitting the data in these ways had essentially no effect on the four estimated cost coefficients. However, some of the nine duration/frequency coefficients moderately differed when estimated on different data subsamples, to an extent that appears not attributable solely to finite-sample statistical variation. As a consequence, some of our WTP estimates remained unchanged, whereas others measurably differed. In sum, we found what appear to be moderate-ordering effects that should give caution to application of some of our estimates. However, the observed effects were neither strong nor pervasive enough to render the estimates unusable. We have no basis to conjecture about the cognitive processes that may generate ordering effects. We think this is an important subject for future research.

4.5. *Willingness to Pay for Reliability.* In this section, we use the parameter estimates in column (3) of Table 3 to estimate the willingness of a consumer with mean preferences to pay for electricity reliability. WTP equals the negative of the ratio of the marginal utility of an outage attribute to the marginal utility of electricity cost. The frequency and duration of outages are negatively valued attributes, so we report the amount of money the household is willing to pay for a reduction in these attributes.

Table 4 evaluates WTP for a unit reduction in F and D at specified values of (F, D) and specified timings for outages. Each column concerns one of the three timing scenarios: weekday peak, weekday off-peak, and weekend off-peak. There are no peak hours on weekends.

In order to enhance comparability of the findings for D and F, the entries in the table divide the WTP for a reduction in D by the number of outages F, and they divide the WTP for a reduction in F by the duration of an outage D. Thus, the table shows the amounts that a person with mean preferences is willing to pay for a one-minute reduction in outage, where the reduction may be achieved by reducing the duration or frequency of outages.

For specificity, Table 4 considers a household with average income. We estimated separate electricity-cost parameters for households with below average and above average incomes, so their WTP are fixed multiples of the calculations presented in the table. Specifically, a household with below (above) average income has 0.51 (3.75) times the WTP amounts shown.[8]

4.5.1. *Findings.* The top panel of the table shows WTP for one-minute reductions achieved by reducing the duration of outages. Consider the findings for weekends. The estimate of 0.98 for $F = 1$ indicates that the consumer is willing to pay \$0.98 for a one-minute reduction achieved by reducing duration when there is one outage per season. We find that WTP declines with F, being \$0.36 per minute when there are five outages per season.

The table shows that WTP during peak hours on weekdays is very similar to that during weekends. In contrast, consumers place little or no value on reductions in the duration of outages during off-peak hours. The estimates are small in magnitude and, for the most part, are statistically indistinguishable from zero.

The bottom panel of Table 4 shows WTP for one-minute reductions achieved by reducing the duration of outages. The estimates for $D = 10$ and $D = 30$ are statistically very imprecise, so we will focus on the much more precise estimates for $D \geq 60$.

Consider the findings for weekends. The estimate of 0.42 for $D = 60$ indicates that the consumer is willing to pay \$0.42 for a one-minute reduction achieved by reducing the frequency of outages when outages have a duration of 60 minutes. WTP declines with D, being 25 U.S. cents per minute when outages have lasted for five hours. The table shows that WTP during weekdays is very similar to that during weekends, both for peak and for off-peak hours. It is noteworthy that most of the WTP estimates are statistically precise, even though the coefficients on F and D and their interactions with the timing variables usually are not.

In sum, the WTP for a one-minute reduction in outage time is similar during weekends and during peak hours on weekdays. Consumers, however, value differently how this one-minute reduction in outage time is achieved. They are willing to pay more when a one-minute reduction in outage time is obtained by reducing the duration of outages than by reducing the number of outages during the season. During the weekday off-peak hours, this pattern is reversed, as consumers are willing to pay somewhat more for reductions in the number of outages than for shorter durations per outage.

The welfare cost of outages is often expressed in terms of money per KWh of electricity unsupplied. In order to translate the estimated WTP amounts of Table 4 into this measure, we multiply them by 60 minutes and divide by the average KWh of electricity consumption, which varies by timing scenario and income group.[9] Table 5 presents these estimates in U.S. dollars per KWh unsupplied.

Our estimates are somewhat higher than those estimated by Beenstock et al. (1998) for the Israeli economy in 1991–1992. The estimates are more in line with those reported for the United States by Doane et al. (1988) and Woo et al. (1991). Comparison of our estimates with earlier ones by other authors should be made with caution because our methodology differs from that of previous work. Whereas the earlier estimates are based on analysis of stated choices, ours is based on analysis of stated-choice probabilities. Comparisons should also be made with caution because the time periods differ and because there is no inherent reason why WTP should be the same in different countries and/or in different periods.

[8] The standard errors of the estimates also change, but not necessarily proportionally.

[9] We use the actual KWh consumed for the households in the sample, which is an average for the whole bi-monthly billing cycle. We adjust this average KWh to the different scenarios using the national averages for electricity consumption in the different scenarios and their relative frequency during the billing cycle.

TABLE 5
MEAN OUTAGE COST (US DOLLARS PER KWH UNSUPPLIED)*

Reduction in D	(1) Weekday, Peak	(2) Weekday, Off-Peak	(3) Weekend, Off-Peak
$F = 1$	42.37	−7.38	47.22
	(23.58)	(26.80)	(23.12)
$F = 2$	28.38	3.26	28.56
	(11.85)	(10.92)	(11.54)
$F = 3$	23.71	6.81	22.34
	(8.35)	(6.22)	(7.95)
$F = 4$	21.38	8.58	19.24
	(6.87)	(4.55)	(6.35)
$F = 5$	19.98	9.65	17.37
	(6.14)	(4.13)	(5.52)
Reduction in F	Weekday, Peak	Weekday, Off-Peak	Weekend, Off-Peak
$D = 10$	−11.56	22.39	72.37
	(51.77)	(92.93)	(54.31)
$D = 30$	5.74	16.73	30.73
	(14.82)	(27.28)	(16.97)
$D = 60$	10.06	15.32	20.32
	(6.19)	(11.35)	(8.07)
$D = 90$	11.50	14.85	16.85
	(4.05)	(6.64)	(5.50)
$D = 120$	12.22	14.61	15.11
	(3.57)	(4.95)	(4.49)
$D = 150$	12.66	14.47	14.07
	(3.58)	(4.46)	(4.07)
$D = 180$	12.94	14.38	13.38
	(3.72)	(4.46)	(3.88)
$D = 300$	13.52	14.19	11.99
	(4.26)	(5.20)	(3.81)

*Cluster-bootstrapped standard errors are in parentheses.

4.6. *The Dispersion of Preferences.* Recall that we assumed a random-coefficient model for utility, with $\beta_i = b + \eta_i$. Subsections 4.4 and 4.5 analyzed the mean preferences b and the derived WTP of persons with these preferences. In this section, we study the dispersion of preferences. Our analysis is quite simple. It does not require parametric distributional assumptions on preferences. Nor does it employ the cumbersome maximum likelihood methods typically applied to random-coefficients models estimated with actual or stated-choice data.

Consider sample member i. We first apply our estimate of b to Equation (6) to estimate the unobserved utility components u_{i2}. We then estimate the person-specific preference parameters η_i by solving the equations $u_{i2} = (x_{i2} - x_{i1})\eta_i$.

We can solve exactly for η when the number of observations (games) per person is equal to the number of regressors in the model (the dimension of η). When the number of observations per person is larger than the number of regressors, we compute the least squares solution for η.

We use the model in column (3) of Table 3 to compute $\hat{u}_{ij}$. Our computations allow the valuation of outages to vary with their timing but assume that timing affects only the mean preference b and not the individual component η_i. Although somewhat restrictive, this assumption greatly increases the number of observations available to estimate η_i. Under this assumption, the model has five random components. Hence, we can compute η for the 448 sample members who played five or more games.

Table 6 presents selected quantiles of the estimated distribution of β_i. The findings have two main features that warrant attention. First, we find considerable dispersion of preferences across the population. In each case, the interquartile range of the preference parameter is large relative

TABLE 6
DISTRIBUTION OF INDIVIDUAL β'S

	10%	25%	50%	75%	90%	N
Weekday, off-peak						
β_F	−2.3389	−0.504	−0.0903	0.2409	1.1457	448
β_D	−0.0889	−0.0171	−0.0020	0.0114	0.0389	448
β_T	−0.0100	−0.0036	−0.0003	0.0024	0.0126	448
Weekday, peak						
β_F	−2.3156	−0.4807	−0.067	0.2642	1.1690	448
β_D	−0.0921	−0.0203	−0.0052	0.0082	0.0357	448
β_T	−0.0102	−0.0038	−0.0005	0.0022	0.0124	448
Weekend, off-peak						
β_F	−2.3784	−0.5435	−0.1297	0.2014	1.1062	448
β_D	−0.0927	−0.0210	−0.0058	0.0076	0.0350	448
β_T	−0.0099	−0.0035	−0.0002	0.0025	0.0127	448
β_C (low income)	−0.0346	−0.0066	−0.0011	0.002	0.0076	156
β_C (average income)	−0.0171	−0.0086	−0.0017	0.0002	0.0027	97
β_C (high income)	−0.019	−0.0065	−0.0015	0.0011	0.0071	116
β_C (missing income)	−0.0501	−0.0091	−0.0024	−0.0002	0.0068	79

TABLE 7
PERCENTAGE OF NEGATIVE COEFFICIENTS

	Weekdays Off-Peak	Weekdays Peak	Weekends Off-Peak	
β_F	59	56	62	
β_D	55	62	53	
β_T	54	54	67	
	Below Average	Average	Above Average	Did Not Respond
β_C	62	70	65	77

to its mean value. Second, although the majority of the estimates of all coefficients are negative as expected, sizable fractions are positive (see Table 7). Positive values are contrary to standard consumer theory, which presumes that households should negatively value electricity outages and price.

Some of the dispersion and positivity of the estimated coefficients may be a consequence of the probability rounding discussed in Section 3 and Subsection 4.3. Recall that the LAD estimation we used to estimate mean preferences in the population is robust to rounding. However, there is no robust way to estimate individual preference parameters. Our use of Equation (6) to estimate person-specific β_i inevitably rests on the rounded probability values that subjects report and, in particular, on the way that we interpret reports of zero probability.

Rounding may partially explain our findings on the distribution of individual preferences, but we doubt that this is the entire story. Our utility model, as any such model, can no more than approximately describe consumer behavior. The model delivers sensible findings regarding mean preferences, but more questionable ones about the full distribution of preferences. This suggests that the model should not be taken as the final word when analyzing consumer preferences for electricity reliability.

4.7. *Maximum Score Estimation of Mean Willingness-to-Pay.* The empirical findings presented in Subsections 4.4–4.6 are based on the mixed-logit model developed in Subsections 3.2 and 3.3. As discussed earlier, this model has many appealing properties. However, it makes the hard-to-motivate assumption that each respondent i believes $(\varepsilon_{i1}, \varepsilon_{i2})$ to be i.i.d. with the extreme value distribution. In this section, we present findings based on the model with weaker assumptions discussed in Subsection 3.4. Here we assume only that each person i places subjective median zero on $\varepsilon_{ij} - \varepsilon_{ik}$, and we apply the maximum score method.

Define y_{ir} for each game $r = 1, \ldots, R_i$ played by individual i to be

$$y_{ir} = \begin{cases} 1 & \text{if } q_{i2}^r \geq 0.5, \\ 0 & \text{otherwise,} \end{cases}$$

where q_{i2}^r is the elicited probability of choosing alternative 2 in game r. The score function can be written

$$S(b) = \sum_{i=1}^{N} R_i - \sum_{i=1}^{N} \sum_{r=1}^{R_i} |y_{ir} - I\{(x_{i2r} - x_{i1r})b \geq 0\}|, \tag{12}$$

where $I\{\cdot\}$ is the indicator function taking the value one when the expression within the curly brackets is true and zero otherwise. Thus, one loses a point when the value of y differs from that of the indicator function.

The maximum score estimate is the set of values of b that minimize the number of wrong predictions, which is

$$S^*(b) = \sum_{i=1}^{N} \sum_{r=1}^{R_i} |y_{ir} - I\{(x_{i2r} - x_{i1r})b \geq 0\}|. \tag{13}$$

Observe that $S^*(b)$ is a step function. Hence, standard derivative-based local optimization routines cannot be applied. Instead, we perform a grid search, which guarantees that we find the global minimum of $S^*(b)$. We first use a coarse grid to locate a neighborhood of the global minimum and then search a finer grid within this neighborhood. We do not report confidence intervals or other measures of precision because asymptotic distribution theory for MS estimation of partially identified models is not available. We only know that the estimates are consistent under the maintained assumptions.

In order to ease the computational burden, we estimate simpler utility function specifications than those underlying the estimates in Table 3. Specifically, we estimate parameters for two specifications that eliminate all interaction terms. These are

$$U_{ij}^1 = \beta_{iC}^1 C_j + \beta_{iT}^1 T_j + \varepsilon_{ij}^1 \tag{14a}$$

$$U_{ij}^2 = \beta_{iC}^2 C_j + \beta_{iD}^2 D_j + \beta_{iF}^2 F_j + \varepsilon_{ij}^2. \tag{14b}$$

In each case, the mean values of the parameters can at most be identified up to scale. In order to fix the scale, we impose the normalization $b_C = -1$. Hence, there is one parameter to estimate in the first specification of the utility function, b_T^1, and two parameters in the second one, b_D^2 and b_F^2.

Consider the first specification. We found that the minimum value of $S^*(b)$ occurs in the short interval $[-1.2772, -1.2223]$. For comparison, LAD estimation of the mixed-logit model yields $b_C^1 = -0.0007817$ and $b_T^1 = -0.0010206$, which imply $\frac{b_T}{b_C} = 1.31$. Thus, the LAD and the MS estimates for $\frac{b_T}{b_C}$ are very similar.

Consider the second specification. We found that the minimum value of $S^*(b)$ occurs in the two small regions $(-5.90, [-116.58, -116.34]) \cup (-5.89, [-116.66, -116.65])$. Thus, the estimate of the first parameter is essentially a point (up to 0.01 rounding) and the estimate of the second parameter is a very short interval. For comparison, LAD estimation of the mixed-logit model yields $b_C^2 = -0.000597$, $b_D^2 = -0.0048388$, and $b_F^2 = -0.1060176$, implying $\frac{b_D}{b_C} = 8.131$ and

$\frac{b_F}{b_C} = 177.6$. Thus, the LAD estimates of $\frac{b_D}{b_C}$ and $\frac{b_F}{b_C}$ are moderately larger than the MS estimates, but not very different.

We are encouraged that the WTP estimates obtained by LAD and MS estimation of simple utility functions specifications are reasonably close to one another. It would be useful to compute MS estimates of richer utility specifications. However, MS estimation of long parameter vectors, such as those estimated by LAD in Table 3, is computationally very difficult. Hence, we suffice with the simpler specifications considered here.

5. CONCLUSION

This article makes two contributions. First, it extends the Manski (1999) approach to the estimation of random utility models with random coefficients using elicited choice probabilities instead of stated choices. The linear mixed-logit model developed in Section 3 requires weaker parametric restrictions than conventional mixed-logit analysis of stated choices and is much easier to implement. However, this model still requires certain distributional assumptions, namely, symmetry of the cross-sectional distribution of preferences and the i.i.d. extreme-value assumption on the subjective distribution of ε. The linear model estimable by maximum score weakens the latter assumption considerably.

Second, the article reports the first application of the "elicited choice probability" methodology. We estimate preferences for electricity reliability in Israel. We use the estimated mean preferences to estimate WTP for reductions in the duration of electricity outages. We find that households with mean preferences are willing to pay significant amounts of money for electricity reliability. During weekends and peak weekday hours, consumers are willing to pay more for a one-minute reduction in outage time obtained by reducing average duration of outages than by reducing the number of outages. However, during the weekend off-peak hours, this pattern is reversed. The methodology allows for estimation of individual preferences in a straightforward manner, and we find that preferences for electricity reliability exhibit considerable dispersion in the population.

As described in the article, eliciting choice probabilities from a sample of households proved to be no more difficult than eliciting stated choices. Choice probabilities allow consumers to express uncertainty about their actual behavior whereas stated choices do not. Thus, it seems that we can get "more for the buck" by adopting this approach in applied research.

APPENDIX

TABLE A.1

LAD ESTIMATES OF UTILITY FUNCTION PARAMETERS FOR SUBSAMPLES

		By Age (years)		By Education		By Income			
Dependent Variable: Log Probability Ratio	Full Sample	50 or Older	Less than 50	College	Non-College	Below Average	Average	Above Average	Non-response
Cost	–	–	–	–	–	−0.00110	−0.000915	−0.000436	−0.00103
						(0.00025)	(0.00027)	(0.00027)	(0.00039)
Cost: (baseline)	−0.00148	−0.00151	−0.00137	−0.00107	−0.00167	–	–	–	–
	[0.00032]	(0.000470)	(0.000371)	(0.000750)	(0.000349)				
Dummy for average income	0.00073	0.000889	0.000561	0.000341	0.00102	–	–	–	–
	[0.00036]	(0.000570)	(0.000448)	(0.000862)	(0.000418)				
Dummy for above average income	0.00128	0.00212	0.000858	0.000624	0.00200	–	–	–	–
	[0.00038]	(0.000508)	(0.000440)	(0.000761)	(0.000500)				
Dummy for missing income	0.00048	0.0000362	0.000837	0.000559	0.000467	–	–	–	–
	[0.00046]	(0.000589)	(0.000482)	(0.00100)	(0.000465)				
Duration	0.00122	0.00158	0.000926	−0.000409	0.00239	0.00753	−0.000525	−0.00266	−0.000630
	[0.00189]	(0.00261)	(0.00250)	(0.00369)	(0.00202)	(0.0024)	(0.0035)	(0.0050)	(0.0039)
Frequency	−0.0049	0.00925	−0.0372	−0.110	0.0661	0.152	−0.0864	−0.0993	−0.114
	[0.0565]	(0.0789)	(0.0660)	(0.0969)	(0.0565)	(0.072)	(0.096)	(0.13)	(0.12)
Total outage time ($T = F \times D$)	−0.00080	−0.00100	−0.000548	−0.000143	−0.00124	−0.00192	−0.000509	−0.000321	0.0000644
	[0.00041]	(0.000548)	(0.000489)	(0.000753)	(0.000408)	(0.00046)	(0.00067)	(0.00099)	(0.00080)
Duration × weekend	−0.0039	−0.00542	−0.00209	−0.00174	−0.00477	−0.00657	−0.000476	−0.00828	0.00113
	[0.0022]	(0.00315)	(0.00291)	(0.00463)	(0.00264)	(0.0030)	(0.0058)	(0.0053)	(0.0045)
Duration × peak	−0.0032	−0.00631	0.000157	0.00122	−0.00392	−0.00968	−0.00205	−0.000426	−0.000640
	[0.0024]	(0.00309)	(0.00318)	(0.00470)	(0.00306)	(0.0031)	(0.0045)	(0.0059)	(0.0067)
Frequency × weekend	−0.0395	−0.0692	0.00411	0.124	−0.131	−0.137	−0.0531	−0.0898	0.165
	[0.0646]	(0.0915)	(0.0812)	(0.116)	(0.0718)	(0.097)	(0.15)	(0.14)	(0.14)
Frequency × peak	0.0233	−0.00635	0.0636	0.144	−0.0156	−0.115	0.0181	0.105	0.192
	[0.0676]	(0.0972)	(0.0827)	(0.119)	(0.0786)	(0.093)	(0.13)	(0.15)	(0.17)
T × weekend	0.00009	0.000464	−0.000359	−0.000895	0.000605	0.00104	0.0000999	0.000391	−0.00161
	[0.0005]	(0.000646)	(0.000586)	(0.000910)	(0.000516)	(0.00065)	(0.0011)	(0.0010)	(0.0010)
T × peak	−0.00022	0.000313	−0.000778	−0.00120	0.000134	0.00109	−0.000144	−0.000652	−0.00144
	[0.00049]	(0.000669)	(0.000614)	(0.000928)	(0.000575)	(0.00062)	(0.00091)	(0.0011)	(0.0013)
Observations	3947	1739	2208	1252	2695	1379	842	1034	692

[1]Boostrapped standard errors based on 500 replications are in parentheses.

REFERENCES

BEENSTOCK, M., E. GOLDIN, AND Y. HAITOVSKY, "Response Bias in a Conjoint Analysis of Power Outages," *Energy Economics* 20 (1998), 135–56.

BEGGS, S., S. CARDELL, AND J. HAUSMAN, "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics* 16 (1981), 1–19.

BEN-AKIVA, M., AND T. MORIKAWA, "Estimation of Switching Models from Revealed Preferences and Stated Intentions," *Transportation Research A* 24A (1990), 485–95.

CAI, Y., I. DEILAMI, AND K. TRAIN, "Customer Retention in a Competitive Power Market: Analysis of a 'Double-Bounded Plus Follow-ups' Questionnaire," *The Energy Journal* 19 (1998), 191–215.

CAVES, D., J. HERRIGES, AND R. WINDLE, "Customer Demand for Service Reliability in the Electric Power Industry: A Synthesis of the Outage Cost Literature," *Bulletin of Economic Research* 42 (1990), 79–119.

DELAVANDE, A., "Pill, Patch, or Shot? Subjective Expectations and Birth-Control Choice," *International Economic Review* 49 (2008), 999–1042.

DOANE, M., R. HARTMAN, AND C.-K. WOO, "Households' Perceived Value of Service Reliability: An Analysis of Contingent Valuation Data," *The Energy Journal* 9 (1988), 135–51.

FISCHER, G., AND D. NAGIN, "Random Versus Fixed Coefficient Quantal Choice Models," in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge, MA: MIT Press, 1981), 273–304.

GOETT, A., K. HUDSON, AND K. TRAIN, "Customer Choice among Retail Energy Suppliers: The Willingness-to-Pay for Service Attributes," *The Energy Journal* 21 (2000), 1–28.

HOROWITZ, J., "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica* 60 (1992), 505–31.

JUSTER, T., "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design," *Journal of the American Statistical Association* 61 (1966), 658–96.

LAWTON, L., M. SULLIVAN, K. VAN LIERE, A. KATZ, AND J. ETO, "A Framework and Review of Customer Outage Costs: Integration and Analysis of Electric Utility Outage Cost Surveys," http://eetd.lbl.gov/ea/EMS/EMS_pubs.html, 2003.

LOUVIERE, J., AND G. WOODWORTH, "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research* 20 (1983), 350–67.

MANSKI, C., "Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics* 3 (1975), 205–28.

——, "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator," *Journal of Econometrics* 27 (1985), 313–33.

——, "Identification of Binary Response Models," *Journal of the American Statistical Association* 83 (1988), 729–38.

——, "The Use of Intentions Data to Predict Behavior: A Best Case Analysis," *Journal of the American Statistical Association* 85 (1990), 934–40.

——, "Analysis of Choice Expectations in Incomplete Scenarios," *Journal of Risk and Uncertainty* 19 (1999), 49–66.

——, "Measuring Expectations," *Econometrica* 72 (2004), 1329–76.

——, "Partial Identification of Counterfactual Choice Probabilities," *International Economic Review* 48 (2007), 1393–410.

——, AND F. MOLINARI, "Rounding Probabilistic Expectations in Surveys," *Journal of Business and Economic Statistics* 28 (2010), 219–31.

——, AND I. SALOMON, "The Demand for Teleshopping," *Regional Science and Urban Economics* 17 (1987), 109–21.

MATZKIN, R., "Nonparametric and Distribution-Free Estimation of the Binary Threshold-Crossing and the Binary Choice Models," *Econometrica* 60 (1992), 239–70.

MCFADDEN, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics* (New York: Academic Press, 1974).

——, AND K. TRAIN, "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics* 15 (2000), 447–70.

REVELT, D., AND K. TRAIN, "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *The Review of Economics and Statistics* 80 (1998), 647–57.

WOO, C.-K., R. PUPP, T. FLAIM, AND R. MANGO, "How Much Do Electric Customers Want to Pay for Reliability? New Evidence on an Old Controversy," *Energy Systems and Policy* 15 (1991), 145–59.