Title
Assessing the Cost of Large-Scale Power Outages to Residential Customers

Journal
Risk Analysis, 38(2)

ISSN
0272-4332

Authors
Baik, Sunhee
Davis, Alexander L
Morgan, M Granger

Publication Date
2018-02-01

DOI
10.1111/risa.12842

Peer reviewed

Risk Analysis, Vol. 38, No. 2, 2018 DOI: 10.1111/risa.12842

Assessing the Cost of Large-Scale Power Outages to Residential Customers

Sunhee Baik,* Alexander L. Davis, and M. Granger Morgan

Residents in developed economies depend heavily on electric services. While distributed resources and a variety of new smart technologies can increase the reliability of that service, adopting them involves costs, necessitating tradeoffs between cost and reliability. An important input to making such tradeoffs is an estimate of the value customers place on reliable electric services. We develop an elicitation framework that helps individuals think systematically about the value they attach to reliable electric service. Our approach employs a detailed and realistic blackout scenario, full or partial (20 A) backup service, questions about willingness to pay (WTP) using a multiple bounded discrete choice method, information regarding inconveniences and economic losses, and checks for bias and consistency. We applied this method to a convenience sample of residents in Allegheny County, Pennsylvania, finding that respondents valued a kWh for backup services they assessed to be high priority more than services that were seen as low priority ($0.75/kWh vs. $0.51/kWh). As more information about the consequences of a blackout was provided, this difference increased ($1.2/kWh vs. $0.35/kWh), and respondents' uncertainty about the backup services decreased (Full: $11 to $9.0, Partial: $13 to $11). There was no evidence that the respondents were anchored by their previous WTP statements, but they demonstrated only weak scope sensitivity. In sum, the consumer surplus associated with providing a partial electric backup service during a blackout may justify the costs of such service, but measurement of that surplus depends on the public having accurate information about blackouts and their consequences.

KEY WORDS: Backup during power outages; electric service reliability; prolonged blackouts

1. INTRODUCTION

Most causes of power outages, such as lightning strikes, falling trees, squirrel electrocutions, or vehicles crashing into poles, cause little prolonged disruption to daily life. These events result in short-term power outages, as evidenced by the median power outage in the United States lasting less than three hours in 2014.(1) On the other hand, widespread and long-lasting outages can have severe individual and societal impacts. Examples include the ice storm that hit southern Québec, Ontario, and northern New York in 1998, leaving many customers without power for several weeks in the dead of winter (affecting 2.3 million people, economic losses of over $4 billion, and the loss of 44 lives),(2) and the extensive outages along the East Coast after Hurricane Sandy (affecting more than 8 million people, economic losses of over $50 billion, and at least 147 direct deaths).(2) These large outages are not limited to extreme weather events, but can also result from a large solar mass ejection (for example, the geomagnetic storm on the United States and Québec power grids that caused a blackout in 1989),(2) as well as physical and cyber attacks on grid infrastructure.(3)

Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA.

*Address correspondence to Sunhee Baik, Department of Engineering and Public Policy, Carnegie Mellon University, Baker Hall 129, 5000 Forbes Avenue., Pittsburgh, PA 15213, USA; tel: (412)-680-0412; sunheeb@andrew.cmu.edu.

While preventing blackouts altogether is too costly for most service territories,(4) new technologies make it possible to sustain critical social services and serve high priority (HP) customer loads during an extended blackout, for example, by islanding distribution feeders using distributed generation, distribution automation, and smart meters.(5) However, these technologies require incremental investment by utilities, and have benefits that are uncertain and difficult to quantify. For this reason, an understanding of the value people place on the services lost during these events is essential for sound decision making.

For years, power companies in the United States have conducted surveys to assess the value that customers attach to reliable electric services. Such estimates are hardly needed for short blackouts affecting industrial and commercial firms, where lost work hours multiplied by hourly revenue (corrected for postoutage rebound) roughly approximates the outage cost. On the other hand, the soft costs experienced by residential households (e.g., not being able to use the air conditioner or run the refrigerator) are more difficult to quantify. Various surveys have asked people about their willingness to pay (WTP) to avoid a hypothetical outage after providing a brief description of an outage and its duration (see the online appendix for an overview). These studies often leave respondents guessing about what a hypothetical outage would entail, providing little detail about the blackout, its geographical extent, the services that would be available and unavailable, and inconveniences and economic losses they might suffer.(6) It is also not trivial to understand the value of having a small amount of power that could serve peoples' HP loads (e.g., lights or air conditioning during summer), compared to full power that also supports somewhat lower priority (LP) loads (e.g., using a speaker dock, DVD/video player, and LED TV to play a game). Most importantly, past studies of residential customers have only asked respondents about outages that last a few hours, providing little information relevant for investment decisions that would minimize the impact of large-scale outages of long duration.

In this article, we develop and demonstrate an elicitation framework to obtain the informed judgments of residential customers about their WTP for full and partial backup service in the event of an extended outage. We illustrate the method with a study of respondents' valuations of a hypothetical 24-hour power outage on a hot summer weekend in western Pennsylvania. In the study, we test the following two hypotheses:

- **H1:** Providing respondents with detailed information about the circumstances of an outage and helping them think through the costs they are likely to experience will lead to more consistent and less uncertain assessments of the value of backup services;
- **H2:** Respondents will value the first 20 A of service to meet their HP demands much more than they value service to meet LP demands (>20 A).

We focus only on service for individuals, but the approach can be generalized to many other outage scenarios, including how people value providing service to others in their communities and to support critical social services (emergency services, food stores, gas stations, etc.).

2. METHODS

2.1. Overview of the Survey Design

Our elicitation procedure was designed to help residential customers think carefully about a specific large-scale outage and systematically reflect on how much they value their full and partial backup service during that outage.(7) The approach helps respondents understand what services would and would not be available in their homes and communities, their personal load profiles as a function of time of day (under normal circumstances or with the full backup service), HP domestic loads they could operate with the partial backup service (under limited availability), and economic losses they might suffer. The framework also allows respondents to express uncertainty in their preferences, and incorporates consistency and bias checks to determine the reliability of responses. Fig. 1 summarizes the design of our elicitation approach.

In the introduction to the survey, we asked respondents to imagine that a large regional blackout occurred on a hot summer weekend as a result of severe weather events in the Midwest (Fig. 2; see the online appendix for the full blackout scenario). Although there was an outage, Pittsburgh's power system was not directly damaged, so power would be restored in 24 hours. Full and partial backup service were then described to respondents, where the full backup service would provide all the electric power

Fig. 1. Overview of the elicitation design, showing the ordering of information provided, exercises that respondents completed, and the timing of willingness to pay (WTP) questions.

Fig. 2. The hypothetical blackout scenario. We told respondents that there were several tornadoes (left) that struck big power lines in the Midwest (right), and resulted in a large regional blackout that spread to the entire Mid-Atlantic and North Eastern parts of the United States (middle).

respondents would normally have used, while the partial backup service would provide only 20 A service for the entire house.

After introducing the scenario and backup services, we elicited respondents' WTP for the full and partial backup service using a multiple bounded discrete choice method, an approach that provides a range of bids that respondents are asked to accept or reject.[8,9] Fig. 3 shows the WTP question used in the study. The range of values from \$0 to \$75 was chosen based on the range of results from a pilot study, and a "not sure" column was included to allow respondents to express uncertainty about their WTP.[10] For each question, respondents indicated their maximum "sure" WTP (the upper limit from the "yes" column) and maximum "not sure" WTP (the upper limit from the "not sure" column). If a respondent had a very high WTP and marked the entire "yes" column, we asked a follow-up question: "What is the largest amount you would be willing to pay to receive the service?"

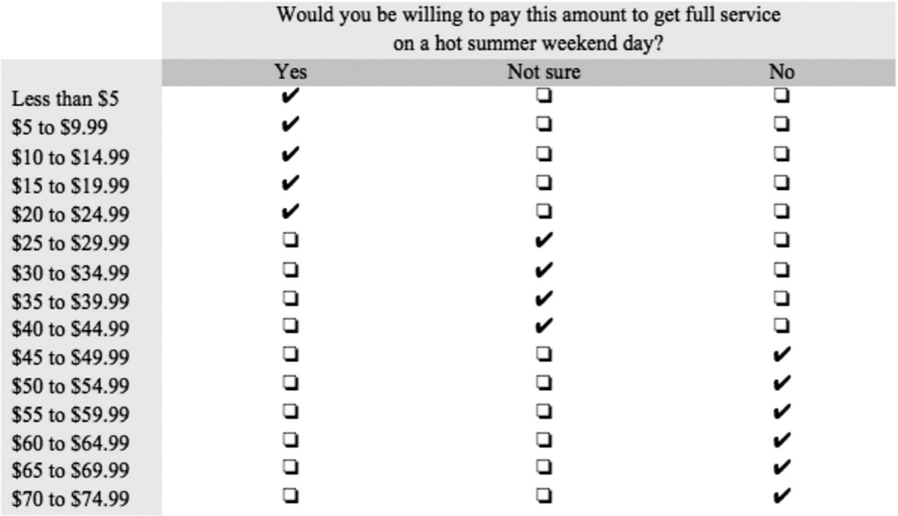
Following this initial WTP assessment, we provided information describing the services that would and would not be available in respondents' homes and communities during the blackout. For example, Table I shows that battery-powered radios and emergency services (including 911) would be available during the blackout, but electric appliances that do not run on batteries, as well as most stores and restaurants without backup generators, would not operate during the blackout.

Next, we asked respondents to play a card stacking game that helped them construct their daily load profiles under normal and limited conditions (in this case, 20 A).[1] Respondents were given a set of cards corresponding to common household appliances. The height of each card was proportional to the amount of power used by that appliance. For example, a typical microwave oven consumes 1,500 watts or 12.5 A at 120V, so the height of microwave oven card was 12.5 cm (left side of the Fig. 4(a)). We divided the day into morning, mid-day, evening, and night, and asked respondents to select the appliances they would likely use in each time period. The height of each stacked column represents the maximum electricity consumed in each period if all appliances are used at the same time (right side of the Fig. 4(a)).[2]

---

[1] Before this study, we considered a number of electric appliance combinations. Based on the results, we chose 20 A as the amount of electricity needed to cover bare necessities.

[2] In some cases, not all appliances would be used at the same time; so, this method provides an upper bound on load. Dealing



	Would you be willing to pay this amount to get full service on a hot summer weekend day?		
	Yes	Not sure	No
Less than $5	✓	☐	☐
$5 to $9.99	✓	☐	☐
$10 to $14.99	✓	☐	☐
$15 to $19.99	✓	☐	☐
$20 to $24.99	✓	☐	☐
$25 to $29.99	☐	✓	☐
$30 to $34.99	☐	✓	☐
$35 to $39.99	☐	✓	☐
$40 to $44.99	☐	✓	☐
$45 to $49.99	☐	☐	✓
$50 to $54.99	☐	☐	✓
$55 to $59.99	☐	☐	✓
$60 to $64.99	☐	☐	✓
$65 to $69.99	☐	☐	✓
$70 to $74.99	☐	☐	✓

Fig. 3. Example response format used in eliciting respondents WTP. In this example, the respondent indicates that he or she would surely pay at least $25, and might be willing to pay as much as $45 for the full backup service during the blackout.

Table I. List Provided to Respondents of Services that Will and Will Not Work in Homes and Communities When the Power is Out for the Entire Region

In Your Home		In Community	
Will Work	Will Not Work	Will Work	Will Not Work
• Old style telephones that have a rotary dial. • Anything that runs on a battery, as long as the battery lasts (e.g., radios, flashlights, laptop computers, and cell phones). • Natural gas and all normal water and sewer services.	• New style telephones that include a plug to a power outlet. • All electrical appliances that cannot also run on batteries, including air conditioners and blowers that circulate air. • Cable and Internet service.	• Emergency service including 911 (via cell phone or rotary dial phone). • Hospitals, police stations, and other places that have backup generators. • TV and radio stations (most have backup generators). • Natural gas and all normal water and sewer services. • Bus service. • GPS service.	• Traffic signals. • Street lights. • Banks and ATMs. • Most gas stations (pumps need electricity). • Food stores (lights, refrigeration, and cash registers will not work). • Most restaurants (very few have backup generators). • Elevators in buildings without backup. • Ventilator fans and lighting in traffic tunnels. • Electric trolley service. • Airport major delays.

Once respondents created their normal load profiles, they were then asked to select a set of HP appliances from their stacked columns to fit under the 20 A limit (Fig. 4(b)). Upon finishing the game, respondents were asked a second time for their WTP for both the full and partial backup service.

Finally, we asked respondents to think about the monetary losses that they would incur as a result of the 24-hour blackout. To do this, we reproduced a recommendation from the U.S. Department of Agriculture regarding perishable foods in refrigerators, and asked respondents to estimate the value of perishable food they have in their refrigerators and

---

with the possible time sequence of appliance usages would have added a great deal of complication, without yielding significant additional insight. We did not mention this issue and most respondents did not bring it up. We wanted respondents to focus on the loads they considered most important, especially when they initially did not understand the concept.

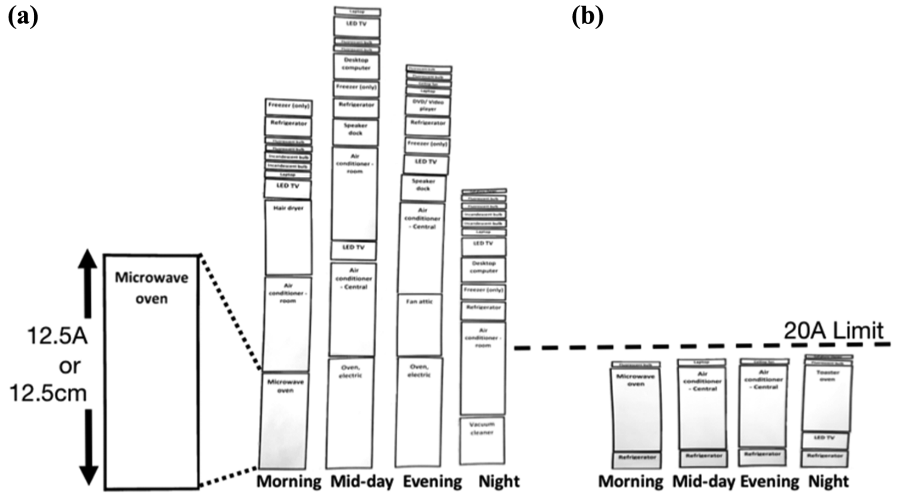


Fig. 4. The electric appliance card stacking game. (a) The height of each card for an appliance or device is proportional to the power consumed, and each respondent built his or her normal electricity consumption profiles for four time periods by using the appliance cards; (b) Each respondent selected his or her high priority (HP) loads to fit under the 20 A limit.

would likely lose in the 24-hour outage. This exercise was followed with a third and final set of WTP questions, again asking respondents to evaluate their WTP for the full and partial backup service.

2.2. Assessment of Bias and Consistency

In addition to providing the information needed to assess respondents' value of the backup services, we tested two important effects that have cast doubt on WTP numbers from contingent valuation studies: scope insensitivity and anchoring.(11) Respondents are scope insensitive if their valuations of a given good or service do not reflect its magnitude. For example, Desvousges *et al.* report that people assigned very similar values (~$80) to protect 2,000, 20,000, and 200,000 birds from being killed by oil spills, suggesting that they cared about protecting the birds, but did not have a precise dollar per bird value in their minds.(12) Anchoring bias occurs when WTP estimates are influenced by irrelevant numerical information. For example, in a classic study, respondents gave higher estimates of the percentage of African countries in the United Nations after they were provided an arbitrary high number in an unrelated task, compared to respondents who were provided with an arbitrary low number.(13)

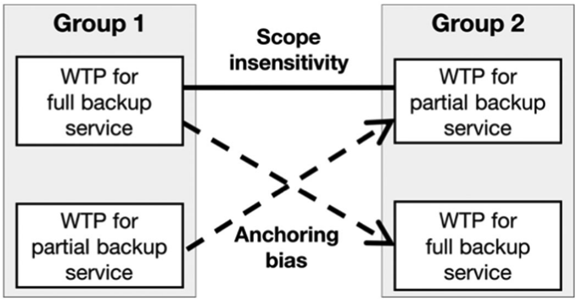


Fig. 5. Experimental design. Group 1 completed the WTP question for the full backup service, then completed the WTP question for the partial backup service. Group 2 completed the WTP questions in the reverse order. Scope insensitivity (solid line) predicts that the WTP for the full backup service in Group 1 should not be different from the WTP for the partial backup service for Group 2. Anchoring (dashed arrow) predicts that the WTP for the full backup service for Group 1 is greater than the WTP for the full backup service for Group 2 (which would be anchored by the lower number preceding it). Similarly, anchoring predicts that the WTP for the partial backup service for Group 1 (with a larger number preceding it) is greater than the WTP for the partial backup service for Group 2.

To test for scope insensitivity and anchoring, we used a 2 × 2 between-subjects design with repeated measures on the second factor, as shown in Fig. 5. Respondents were randomly assigned to Group 1 or Group 2 by a virtual computer-generated coin toss (http://www.random.org). Group 1 first gave their WTP for the full backup service, and then moved on to the partial backup service, whereas Group 2 responded in the reverse order. If respondents are scope insensitive, Group 1's initial WTP for the full backup service and Group 2's initial WTP for the partial backup service should not differ, suggesting respondents care about getting a service, but do not have a specific *dollar-per-amp* (or *dollar-per-kWh*) figure in their minds. If respondents are biased by anchoring, Group 2's WTP for the full backup service should be smaller than Group 1's, as the partial service WTP question for Group 2 anchors respondents on a lower number for their full service WTP. Using the same reasoning, Group 1's WTP for the partial backup service should be greater than Group 2's WTP for the partial backup service.

We also developed two additional conditions to check the consistency of respondents' preferences. Our first check was whether WTP for electricity backup per kWh were greater than or equal to the normal electricity cost (i.e., \$0.11/kWh), as the value of electricity should not be decreased by a blackout.[3] Second, for the same respondent (as opposed to across experimental groups), the WTP for the full backup service should be greater than or equal to the partial backup service, as the former encompasses the latter.

3. RESULTS

To recruit a diverse sample within Allegheny County, the study was advertised through local community organizations and online through Craigslist and the Center for Behavioral Decision Research at Carnegie Mellon University. Individuals were required to be 25 years or older, had to have lived in Allegheny County for at least three years, and have at least one other adult living in their household.[4] All interviews were conducted in a face-to-face format between July and August 2015. The respondents completed the three sections of the survey individually at their own pace. Interviews took one hour on average. Once the interview was completed, the respondents were compensated \$10 for their time. We recruited 73 eligible respondents (Group 1: $n = 38$, Group 2: $n = 35$). We excluded three interviews because one respondent did not meet the eligibility criteria, one already owned a number of backup generators, and one could not understand the WTP response mode. The conclusions do not change if we include the first outlier (WTP results were similar to the averages), and no results could be calculated for the other two.[5,6]

We compared the demographic information of the survey respondents with census data for Allegheny County, Pennsylvania. The survey sample was similar with respect to income and race, but had fewer men and middle-aged individuals than the local population. The average age of the respondents was 43 ($SD = 16$), 56% were female, and 33% were nonwhite. On average, the respondents had lived in Allegheny County for 20 years ($SD = 18$). About 73% lived in the Greater Pittsburgh metropolitan area, and 27% of the respondents lived in the suburbs.

3.1. The Value of Service for Loads of HP and LP

Our first result is that the respondents valued backup service for their HP loads more than that for their LP loads, and as they received more information, their WTP for the partial backup service to meet HP loads ($\leq$20 A) increased, while their WTP for power to serve LP loads ($>$20 A) decreased slightly.

We calculated the amount that the respondents were sure they were willing to pay to meet HP demands in the following way: (1) we used the upper bound of the highest box the respondents checked in the "yes" column of the WTP question for the partial service, (2) estimated the respondents' electricity consumption by summing up the product of the amount of electricity that each appliance consumes, the number of each selected appliances, and the time that each appliance would be turned on,[7]

---

[3] In the elicitation, we did not tell the respondents about the normal price of electricity to avoid anchoring them on a value we provided.

[4] In other words, we recruited residents of Allegheny County who have a sense of their domestic budget and experience paying their electric bills, have lived long enough in, and are familiar with, the region and its power system, and are electric consumers. The criteria for eligibility were tested in the pilot study, and slightly modified before the actual implementation.

[5] The second outlier was completely off-grid, thus the hypothetical outage scenario and the assumptions for the backup services were not applicable (this respondent refused to answer the survey). The third outlier's answers were not at all consistent (e.g., "not sure" WTP was lower than "sure" WTP but higher than "no" WTP), so we could not calculate the value to serve HP and LP loads and the respondent's range of uncertainty.

[6] In addition to the main results from the analyses of the respondents' WTP, we performed additional analyses to compare the respondents' WTP by demographic category (household income and housing types) and level of preparedness for an outage (whether they had backup generators, battery power devices, etc.). We found that the respondents' WTP were slightly influenced by their income levels, but not by other variables. We also compared the respondents' WTP in relation to their outage experiences during their lifetime, but we did not observe any significant difference. Additional details and results are provided in the online appendix.

[7] Because we only asked the respondents to estimate their loads at four specific times of the day, we did not have their actual total

Table II. Summary of the "Sure" Value per kWh ($/kWh) to Serve High Priority (HP) and Lower Priority (LP) Loads

		Min	Median	Mean (*M*)	Max	*SD*	*N*
Initial	HP loads	0.00	0.58	0.75	3.1	0.63	73
	LP loads	0.00	0.29	0.51	8.4	1.0	73
Middle	HP loads	0.00	0.82	1.1	3.9	0.84	73
	LP loads	−0.61	0.17	0.36	5.1	0.79	73
Final	HP loads	0.00	0.92	1.2	5.2	0.88	73
	LP loads	−0.63	0.14	0.35	5.1	0.76	73

and (3) divided the maximum WTP for the partial backup service by the amount of power consumed by the appliances they selected within the 20 A limit.[8] Thus, the value of meeting HP demands would be $\frac{WTP_{partial}}{Electricity\ consumption\ within\ 20\ A\ limit}$. For example, if a respondent indicated his/her maximum willing to pay of $25 for 10 kWh (from the partial service), the value of serving HP demands would be $\frac{\$25}{10} = \$2.5/\text{kWh}$. Next, we calculated the amount that the respondent was sure he/she was willing to pay to serve LP demands by: (1) using the upper bound of the highest box that the respondents checked in the "yes" column of the WTP question for the full service (maximum WTP for the full service), (2) subtracting that number from their maximum WTP for the partial service, and (3) dividing by the amount of power consumed by the appliances they selected without any limit, minus the power consumed by appliances they selected within the limit. Thus, the value of serving LP demands would be $\frac{WTP_{full}-WTP_{partial}}{Elec.consumption_{\text{w/o limit}}-Elec.consumption_{\text{within 20 A limit}}}$. For the example in Fig. 3, because the respondent indicated that he/she was willing to pay up to $45 for 70 kWh (from the full service), his/her value of serving LP demands would be $\frac{\$45-\$25}{70-10} = \$0.33/\text{kWh}$.

Table II summarizes the sure amount the respondents were willing to pay per kWh for serving HP and LP demands, and Fig. 6 shows each observation. We compared the values using the Wilcoxon signed-ranks tests (Wsr).(14) We report the statistic V, which describes the smaller of the sum of positive signed ranks and the sum of negative signed ranks, for the initial WTP (V_i), middle WTP (V_m), and final WTP (V_f), as well as the difference between middle and initial (V_{mi}), and final and middle (V_{fm}).[9] As Fig. 6 shows, the respondents report a significantly higher WTP to serve HP demands than their LP demands at all stages (Wsr, $V_i = 554$, paired Cohen's $D = 0.21$; $V_m = 270$, $D = 0.67$; $V_f = 212$, $D = 0.75$; all $p < 0.05$), and the values to serve HP demands significantly increased as the survey progressed (Wsr, $V_{mi\_HP} = 80$, $D_{mi\_HP} = 0.58$; $V_{fm\_HP} = 137$, $D_{fm\_HP} = 0.21$, both $p < 0.05$). In contrast, the values to serve LP demands significantly decreased from initial to middle assessments (Wsr, $V_{mi\_LP} = 695$, $p_{mi\_LP} < 0.05$, $D_{mi\_LP} = 0.12$), but did not differ between middle and final assessments (Wsr, $V_{fm\_LP} = 297$, $p_{fm\_LP} = 0.77$, $D_{fm\_LP} = 0.026$).

Thus, by the end of the process (as well as at the other stages), on average, the respondents placed a higher value on serving their HP loads (mean [M] = $1.2/kWh) than that of their LP loads (*M* = $0.35/kWh). Furthermore, their WTP to serve HP demands significantly increased by 56% (from $0.75/kWh to $1.2/kWh) as they came to better understand the inconveniences and monetary losses they might suffer. We found that several respondents decreased their WTP as the survey progressed (full: eight or less, partial: nine or less), indicating the respondents felt free to either increase or decrease their WTP.

3.2. Uncertainty in WTP Assessments

We operationalized the respondents' uncertainty about their WTP as the difference between the upper bound of the highest box that the respondents checked in the "not sure" column of the WTP questionnaire, and the upper bound of the highest box that the respondents checked in the "sure" column. For example, in case of Fig. 3, the respondent checked "yes" up to $24.99, "not sure" up to $44.99, and then "no" afterward. Then, the respondent's range of uncertainty would be $44.99 − $24.99

---

consumptions. Instead, we used three different sets of plausible assumptions and computed the average of the three values in order to estimate the total electricity consumptions. However, the ability to purchase capacity during the blackout is also important. For that, we also conducted the same analysis for capacity charge (for kW). See the online appendix for the results.

[8] Sensitivity analyses for the WTP threshold (the lower, middle, or upper bound of the highest box) and the amount of electricity consumed from the partial backup service (the respondents' actual electricity consumption or 20 A), are presented in the online appendix.

[9] The test statistic for the Wsr (V) is defined as the smaller of the sum of the positive ranks (V^+) or the negative ranks (V^-), where the sum of V^+ and V^- equals the sum of all the ranks ($\frac{n(n+1)}{2}$) if no ranks are tied. If the test statistic significantly deviates from the critical value, we rejected the null hypothesis that the two samples were drawn from the same population distribution.(14)

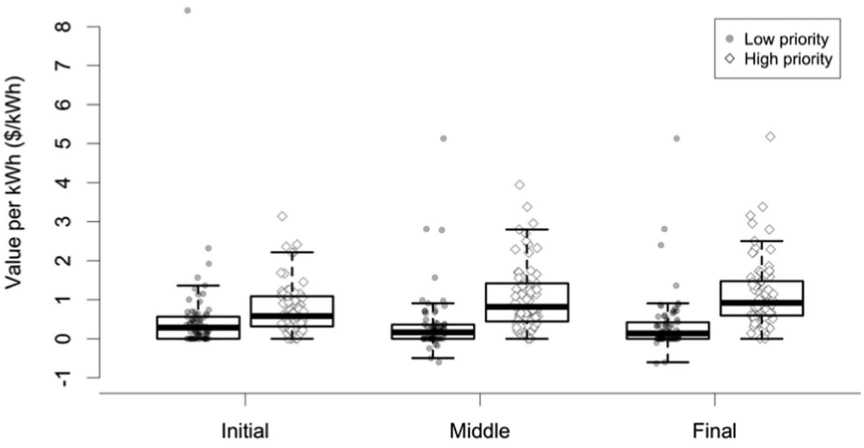


Fig. 6. Distribution of the value per kWh to serve lower priority (LP) and HP loads by stage over the course of the study (LP: left at each stage, HP: right at each stage). Boxplots show the median, interquartile range, and whiskers at 1.5 times the interquartile range (or the greatest/smallest number). Gray circles indicate the value of LP demands for each respondent, and white diamonds indicate the value of HP demands for each respondent at each stage.

= \$20. There were five (partial backup service) and 13 (full backup service) respondents who were willing to pay more than \$75 in the final stage, after which we asked for a single number that best represented their WTP. Because we were not able to obtain a range of uncertainty for these respondents, we excluded them from the uncertainty analysis.

Table III summarizes the results, and Fig. 7 compares the range of uncertainty between different backup services and stages. We also compared the results using the Wsr, and reported the statistic V for the difference between initial and middle (V_{im}), middle and final (V_{mf}), and initial and final (V_{if}) for each backup service, as well as the level of uncertainty in initial (V_i) and final (V_f) stage between the backup services.[10] In the initial stage, the respondents were slightly more uncertain about the partial service than the full backup service (Wsr, $V_i = 237$, $p = 0.44$, paired Cohen's $D = 0.16$). Comparing the initial to middle stage, the respondents became less uncertain, and the decrease was more pronounced in the partial backup service case (Wsr, $V_{im_{full}} = 263$, $p_{im_{full}} = 0.17$, $D_{im_{full}} = 0.19$; $V_{im_{partial}} = 522.5$, $p = 0.06$, $D_{im_{partial}} = 0.21$). Comparing the middle and final stage, the information regarding the respondents' monetary losses slightly decreased their uncertainty about both backup services, but the deceases were not statistically significant (Wsr, $V_{mf_{full}} = 165.5$, $p_{mf_{full}} = 0.66$, $D_{mf_{full}} = 0.082$; $V_{mf_{partial}} = 261$, $p_{mf_{partial}} = 0.34$, paired $D_{mf_{partial}} = 0.14$). Over the course of the entire study, the respondents decreased their uncertainty in their WTP for both the full and partial backup service by 16% and 23%, respectively (Wsr, $V_{if_{full}} = 300.5$, $p_{if_{full}} = 0.07$, $D_{if_{full}} = 0.26$; $V_{if_{partial}} = 628$, $D_{if_{partial}} < 0.05$, $D_{if_{partial}} = 0.28$). Yet, even by the end of the study, uncertainty about the unfamiliar ideas of the partial backup service remained slightly higher than that for the full backup service (Wsr, $V_f = 101$, $p_f = 0.16$, $D_f = 0.18$).

In summary, the information provided by the survey protocol helped respondents better understand the blackout scenario, its consequences, and the backup services. On average, it reduced the range of uncertainty for both the full (\$11 to \$9.0) and partial backup service (\$13 to \$11). The greater uncertainty for the partial versus full backup service likely reflects the respondents' different familiarity with the two options. While the standard deviations and the ranges of uncertainty were fairly high, a primary reason for this was the large heterogeneity across people in their WTP due to different electricity use profiles, demographics, and needs. More discussions about

---

[10] Because the number of respondents with "sure" WTP higher than \$75 for the full and partial backup service in the end of the study did not match (five respondents for the partial backup service and 13 respondents for the full backup service in the final stage), we dropped the respondents with WTP higher than \$75 for the full backup service and compared the ranges of uncertainty. For the comparisons between stages, we dropped WTP higher than \$75 for each backup service in the end of the study (i.e., five respondents from the partial backup service and 13 respondents from the full backup service) to compare the range of uncertainty within each service.

Table III. Summary of the Ranges of Uncertainty After Dropping the Respondents Who Had WTP Higher than $75 in the Final Stage

								Percentages that		
		Number of Respondents	Mean (*M*)	*SD*	Median	Min	Max	Not Sure > Sure	Not Sure = Sure	Not Sure < Sure
Initial	Partial	68	13	13	10	0	55	72%	28%	0%
	Full	60	11	10	10	0	40	70%	30%	0%
Middle	Partial	68	11	12	5.0	0	50	66%	34%	0%
	Full	60	9.6	11	5.0	0	50	67%	33%	0%
Final	Partial	68	11	13	5.0	0	50	59%	41%	0%
	Full	60	9.0	11	5.0	0	50	57%	43%	0%

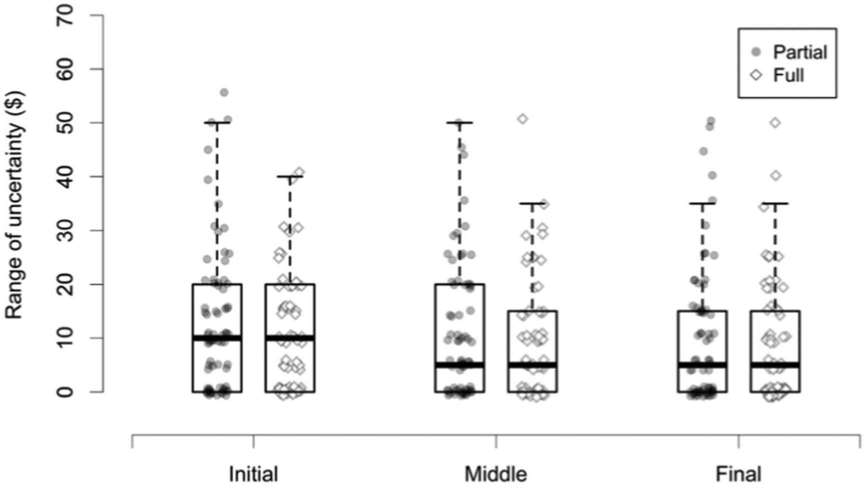


Fig. 7. Distribution of the range of uncertainty in the partial (circles) and full (diamonds) backup service with boxplots after dropping the respondents who had WTP higher than $75 at each stage and backup service (partial: left at each stage, full: right at each). Circles indicate the range of uncertainty for the partial backup service for each respondent, and diamonds indicate the range of uncertainty for the full backup service for each respondent at each stage.

Table IV. Number of Inconsistencies from Two Consistency Checks

		Initial	Middle	Final
Elec. Cost ($/kWh)	Partial	6	3	2
(normal > backup)	Full	8	7	4
WTP (partial) > WTP (full)		0	6	3

heterogeneity and further analysis are provided in Section 3.4.

3.3. Consistency and Bias Checks

We introduced two conditions to check the internal consistency of respondent's WTP assessments. Results are summarized in Table IV. First, we compared the WTP for backup service per kWh with the normal electricity cost (assuming an average electricity rate of $0.11/kWh). At the beginning of the survey, eight respondents valued the full backup service lower than their normal electric services, as did six respondents for the partial backup service. By the end of the survey, these numbers dropped to four (for full) and two (for partial). Second, no respondent gave a higher WTP for the partial backup service than the full backup service at the beginning of the survey, but six did in the middle and three did by the end. Thus, by the end of the survey, 90% of the respondents gave responses that suggested well-reasoned, systematic preferences.

Next, we tested scope sensitivity by comparing the WTP distribution for those first asked to give their partial WTP versus those first asked to give their full WTP in the initial stage. Because this test

Table V. Summary of KS Tests for Anchoring Bias with Summary of Initial WTP Results from Two Groups for the Partial and Full Backup Service

		N	Mean [*M*]	*SD*	*KS-D*	*p*
Group 1 (Full first)	Partial	38	$19	14	0.19	0.49
Group 2 (Partial first)		35	$27	20		
Group 1	Full	38	$35	29	0.22	0.35
Group 2		35	$45	33		

was between subjects, the respondents were not influenced by any prior numerical information when making their judgments. Scope insensitivity implies that there will be little difference between the WTP distributions for these two groups. Although Group 1 gave a higher average value for the full backup service ($M = \$35$, $SD = 29$) than Group 2 gave for the partial backup service ($M = \$27$, $SD = 20$), the results of a two-sample Kolmogorov–Smirnov (KS) between the two groups were not statistically significant ($KS - D_{full_{partial}} = 0.17$, $p_{full_{partial}} = 0.68$). More in-depth analysis about scope insensitivity is provided in Section 3.4.

Next, anchoring bias would be present if the respondents' later WTP estimates are influenced by their earlier estimates. The null hypothesis is that initial WTP distributions of the same backup service were drawn from the same population distribution (H_0: $X1_{initial_j} = X2_{initial_j}$ where $j = full, partial$). We conducted two-sample KS tests and compared Group 1's and 2's cumulative distribution functions, as shown in Table V. In all the cases, the null hypotheses cannot be rejected (two-sample KS test, $KS - D_{full} = 0.22$, $p_{full} = 0.35$; $KS - D_{partial} = 0.19$, $p_{full} = 0.49$). Importantly, Group 1 started with higher WTP (from the full backup service question) but resulted in lower numbers than Group 2, which was the opposite of an anchoring effect. Thus, we conclude that the order of introducing two backup services neither anchored nor influenced the respondents' WTP.

3.4. Multiple Linear Regression Analysis and Scope Sensitivity

We used multiple linear regression to model the respondents' final WTP for the full and partial backup service. To do this, we: (1) included regressors in the model if they reduced the root mean squared prediction error from five-fold cross validation, (2) examined the correlations between the regressors,[11] (3) conducted principal component analyses on the regressors,[12] (4) checked for necessary transformations of the regressors, and (5) modeled the final WTP as linear functions of selected variables and components that minimize the sum of squared residuals (Fig. 8). Both models provide reasonable estimates of respondents' final WTP (adjusted $R^2_{partial} = 0.40$, $p_{partial} < 0.05$ and adjusted $R^2_{full} = 0.38$, $p_{full} < 0.05$) and perform substantially better than simple models only with intercepts (according to five-fold cross-validation). We also used a multilevel model with varying intercepts by respondent (Fig. 9).(15),[13] For more detailed regression results, see the online appendix.

A likely reason for the lack of statistical significance of the scope sensitivity test was the large heterogeneity across people in their WTP (variance of random intercepts across the respondents are 0.64 (partial backup service) and 0.78 (full backup service)). Using the regression models with varying intercepts (Fig. 9, including the outliers), we estimate that heterogeneity, and provide a more precise estimate of scope sensitivity by comparing Group 1's WTP for the full backup service and Group 2's WTP

---

[11] If a regressor is not strongly correlated with other regressors, we did not transform the regressor. However, for some regressors that are strongly correlated with each other, we: (1) scaled (set standard deviation to 1), (2) centered (shifted means to 0), and (3) conducted a principal component analysis to find a linear combination that explains the most variance of the group of variables.

[12] There were seven (partial backup service) and 11 (full backup service) respondents who had an absolute difference between actual and predicted values larger than 1 in the final stage. The two most extreme outliers were the respondents with zero WTP for both backup services (Cook's $D_{partial} = 0.41/0.036$ ($M = 0.021$), $Leverage_{partial} = 0.34/0.10$ ($M = 0.10$); $D_{full} = 0.27/0.24$ ($M = 0.018$), $Leverage_{full} = 0.32/0.10$ ($M = 0.11$)). While removing the two outliers increased some of the variables' coefficients (e.g., value of perishable food), the coefficients of other variables decreased (e.g., electricity consumption under 20 A limitation) because they had almost opposite preferences. The number of responses for the full and partial backup service are different because of some nonresponses in the case of the partial backup service.

[13] There were three/two/two (partial backup service, in the initial/middle/final stages) and one/one/one (full backup service, in the initial/middle/final stages) respondents who had absolute difference between actual and predicted value larger than 1. The two most extreme outliers had zero WTP in the initial stage, but then increased their numbers (to $30 and $50, respectively) in the final stage.

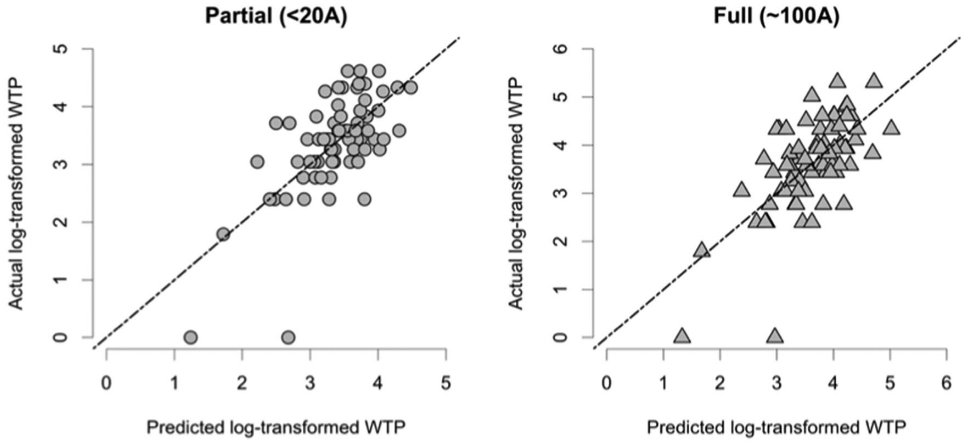


Fig. 8. Predicted log-transformed WTP against actual log-transformed WTP in the final stage (partial: left and full: right) using multiple linear regression analysis against actual final WTP, including two extreme outliers who were not interested in using either backup service.

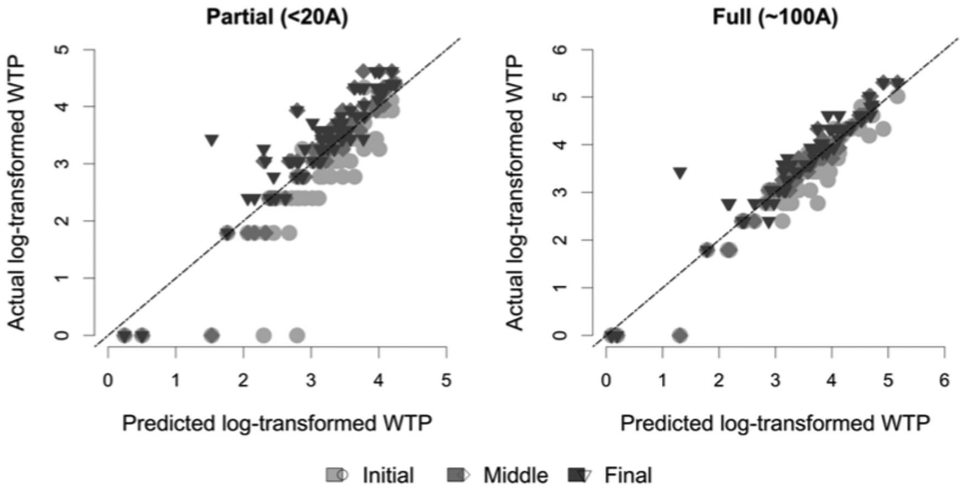


Fig. 9. Predicted log-transformed final WTP against actual log-transformed final WTP (partial: left and full: right) using regression models with varying intercepts by respondents, including two extreme outliers who were not interested in using either backup service.

for the partial backup service. In this model, we included a factor variable for the randomly assigned group (1 = Group 1, 2 = Group 2). If this variable was statistically significant, then respondents would be scope sensitive across their three choices. However, this was not the case ($p \geq 0.37$).

To determine whether we had enough respondents to adequately detect scope sensitivity, we calculated the required sample size needed to reject the null hypothesis of no scope sensitivity 80% of the time with an alpha level of 0.05 by using the effect sizes obtained in the study. According to the result (using a two-sample *t*-test), we would need 226 respondents from each group (mean difference = 0.26, pooled $SD = 1.0$; WTP results were log-transformed before the calculation).

4. DISCUSSION

4.1. Study Results and Policy Implications

First, our results suggest that the value of serving HP demands for a one-time 24-hour outage (*M*= $0.75/kWh) was significantly higher than that of LP demands (*M*= $0.51/kWh) even when they only brought their prior knowledge to the assessment. Second, as the respondents received additional information, they placed higher value on sustaining services they considered HP ($M=$ \$0.75/kWh to \$1.2/kWh), whereas the value they attached to LP demands slightly decreased ($M=$ \$0.51/kWh to \$0.35/kWh). Third, the respondents' uncertainty about their WTP decreased as they worked their way through the protocol (full= \$11 to \$9.0, partial= \$13 to \$11 on average), suggesting that they progressively understood more about the backup services, and how much they cared about those services. Finally, our checks suggested that the vast majority of respondents (90%) were consistent and systematic about their preferences, and were not biased by their previous WTP responses. However, the respondents demonstrated only weak sensitivity to the magnitude of service provided (scope sensitivity).

So far, our study is most similar to Sullivan *et al.*,(16) which combines individual study results conducted by major utilities, and derives a customer damage function for industrial, commercial, and residential customers. The average interruption cost can be estimated as a function of interruption attributes and customer characteristics (see the online appendix for more details).(16) While the results analyzed by Sullivan *et al.* cannot be simply compared to the results obtained in this study (see the online appendix for the reasons), our survey framework has three major improvements with policy implications. First, we can assess the difference between the value for the first few kWh and the last amount consumed.(17) Using the considerable amount of consumer surplus, distribution utilities and other relevant parties, such as regional authorities and local or state governments, can substantially reduce interruption costs if they continue to supply at least a small amount of electricity during such outages. In the event of widespread outages of longer durations, the benefits from implementing a partial backup service and covering customers' bare necessities will become even greater.

Second, the survey framework highlights the benefits of information about outages and associated costs, especially when respondents are not familiar with the issue. If we assume that the cost increases proportionally to the duration, scaling the Sullivan *et al.* results to a 24-hour outage suggests a cost of \$46 (by using the simple linear interpolation: $3.9 + \frac{\$32-\$3.9}{16\,hours-momentary(0\,hour)}(duration) = 3.9 + 1.8(duration)$).(16) While the estimated cost is higher than our study's initial "sure" WTP for the full backup service ($M=$ \$39), it is less than the number we got from the final stage ($M=$ \$51). Importantly, the increase comes from the HP demands (Wsr, $V_{if_{partial}} = 64.5$, $p_{if_{partial}} < 0.05$), not from the LP demands (Wsr, $V_{if_{(full-partial)}} = 462$, $p_{if_{(full-partial)}} = 0.13$). This result suggests that the information did not simply increase the respondents' WTP values, but it actually helped them to better understand which loads were most important to them and why sustaining service to those loads is important.(18) Thus, our study reemphasizes the need for information when eliciting values to serve HP and LP demands.

Finally, our approach demonstrates that respondents expressed a significant amount of uncertainty about their preferences, but part of that uncertainty could be reduced with additional information. Yet, uncertainty persisted throughout the study (full: \$9.0, partial: \$11 on average), illustrating the need for frameworks that can incorporate the uncertainty of public preferences into decision making.

4.2. Limitations

We note three limitations of the study. First, there are some drawbacks associated with the multiple-bounded discrete choice method.(19) For example, Roach *et al.* determine that welfare estimates can be affected by the range of bids available to respondents (range bias),(8) and Alberini *et al.* suggest that order of presentations can have a significant effect.(9) While we alleviated some of the range bias by using the follow-up question, we could not eliminate the range bias entirely, and still had a small peak near \$75 (the maximum for the multiple-bounded discrete choice procedure).[14] Second, we found that the respondents' income levels only slightly influenced their WTP for backup services, suggesting they were either not constrained by their ability to pay, or were not considering the other possible uses of their money.(20) Third, because the length of the interviews precluded our exploring WTP for the partial service levels of other than 20 A, we were not able to trace out the full shape of the consumer surplus.

---

[14] We compared the "sure" WTP distributions without the follow-up question (categorizing all the respondents with WTP higher than \$75 as maximum "sure" WTP is \$75 group) and with the follow-up question (assuming their maximum WTP answers from the follow-up question as their "sure" WTP), and observed reductions in the \$75 peak (65% from the full service and 45% from the partial service in the final stage). Thus, the follow-up question helped alleviate the range bias.

5. CONCLUSION

Low-probability high-consequence interruptions in electric services of large spatial scale and long duration can give rise to enormous economic and social costs, including loss of life.[15] These costs can be reduced if a small supply of electricity can be provided during such outages. However, providing this capability requires incremental investments. One important input to determining whether and where such investments might be warranted is an informed judgment by residential and other customers of the value of such service as reflected through judgments about their WTP. The method we have developed and demonstrated in this article points the way to obtaining such informed judgments.

ACKNOWLEDGMENTS

This work has been supported by the Center for Climate and Energy Decision Making through a cooperative agreement between the National Science Foundation (SES-0949710) and Carnegie Mellon University and by the Thomas Lord Chair and other academic funds from Carnegie Mellon University.

REFERENCES

1. IEEE Benchmark Year 2015 Results for 2014 Data. IEEE Working Group on Distribution Reliability. Available at: http://grouper.ieee.org/groups/td/dist/sd/doc/Benchmarking-Results-2014.pdf, Accessed February 8, 2016.
2. Billion-Dollar Weather and Climate Disasters: Table of Events. National Centers for Environmental Information. Available at: http://www.ncdc.noaa.gov/billions/events, Accessed February 8, 2016.
3. How the EPRI Geo Magnetic Disturbances (GMD) Research Fits Together. Electric Power Research Institute. Available at: http://www.epri.com//abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001026425, Accessed February 27, 2016.
4. Morgan MG. Oral testimony from Prof. M. Granger Morgan prepared for the 2015 September 10 hearing on “Examining the Vulnerabilities of Americas Power Supply” before the Committee of Science, Space and Technology Subcommittee on Oversight and Subcommittee on Energy of the U.S. House of Representatives. Available at: http://docs.house.gov/meetings/SY/SY21/20150910/103914/HHRG-114-SY21-Wstate-MorganM-20150910.pdf, Accessed March 15, 2016.
5. Narayanan A, Morgan MG. Sustaining critical social services during extended regional power blackouts. Risk Analysis, 2012; 32(7):1183–1193.
6. Sullivan MJ, Keane DM. Outage Cost Estimation Guidebook. Electric Power Research Institute, No. EPRI TR-106082, December 1995.
7. Fischhoff, B. Value elicitation: Is there anything in there? American Psychologist, 1991; 46(8):835–847.
8. Roach B, Boyle KJ, Welsh M. Testing bid design effects in multiple-bounded, contingent-valuation questions. Land Economics, 2002; 78(1):121–131.
9. Alberini A, Boyle K, Welsh M. Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. Journal of Environmental Economics and Management, 2003; 45(1):40–62.
10. Cubitt RP, Navarro-Martinez D, Starmer C. On preference imprecision. Journal of Risk and Uncertainty, 2015; 50(1):1–34.
11. Kahneman D, Ritov I, Schkade D. Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. Journal of Risk and Uncertainty, 1999; 19(13):203–235.
12. Desvousges WH, Johnson FR, Dunford RW, Hudson SP, Wilson KN, Boyle KJ. Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy. RTI Press, 1992. Available at: http://www.rti.org/sites/default/files/resources/bk-0001-1009_web.pdf, Accessed on May 24, 2017.
13. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science, 1974; 185(4157):1124–1131.
14. Siegel S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
15. Gelman A, Hill J. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge: Cambridge University Press, December 2006.
16. Sullivan M, Schellenberg J, Blundell M. Updated Value of Service Reliability Estimates for Electric Utility Customers in the United States. Berkeley, CA: Ernest Orlando Lawrence Berkeley National Laboratory, No. LBNL-6941-E, January 2015. Available at: https://eetd.lbl.gov/sites/all/files/lbnl-6941e_0.pdf, Accessed February 8, 2016.
17. Harish SM, Morgan GM, Subrahmanian E. When does unreliable grid supply become unacceptable policy? Costs of power supply and outages in rural India. Energy Policy, 2014; 68:158–169.
18. Orne MT. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962; 17(11):776–783.
19. Cameron TA, Poe GL, Ethier RG, Schulze WD. Alternative non-market value-elicitation methods: Are the underlying preferences the same? Journal of Environmental Economics and Management, 2002; 44(3):391–425.
20. Donaldson C. Valuing the benefits of publicly-provided health care: Does ability to pay preclude the use of willingness to pay? Social Science & Medicine, 1999; 49(4):551–563.
21. Economic Benefits of Increasing Electric Grid Resilience to Weather Outages. Executive Office of the President. August 2013. Available at: https://energy.gov/sites/prod/files/2013/08/f2/Grid%20Resiliency%20Report_FINAL.pdf, Accessed February 9, 2016.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher’s website. Also, the data with source code may be found at https://osf.io/eqp3d/:

Table A1: Sample survey design for residential customer survey.

---

[15] According to the Energy Information Administration, more than 85% of outages to the bulk electric system are caused by severe weather (e.g., thunderstorms, hurricanes, and blizzards); the annual cost of power outages caused by these events is estimated to be $18–$33 billion.(21)

Table C1: Demographic information of respondents in this survey.

Figure C1: Map of where survey respondents live.

Table C2: Longest outage experience from survey respondents.

Table D1: Summary of the "sure" value per amps and kW of high and lower priority loads.

Table D2: Summary of the paired Wilcoxon signed-rank tests comparing "sure" value per amps and kW of high and lower priority loads.

Table E1: Information collected and used in data analysis.

Table E2: Summary of the five-fold cross-validations results for the partial backup service.

Table E3: Summary of the five-fold cross-validations results for the full backup service.

Figure E1: Scatter plots of the selected variables versus respondents' final WTP for the partial backup service.

Figure E2: Scatter plots of the selected variables versus respondents' final WTP for the full backup service.

Figure E3: Correlation matrix plots of selected variables for the partial (above) and full (below) backup service.

Table E3: Summary of principal components and selected variables for the partial (left) and full (right) backup service.

Figure E4: Scatterplots with the final WTP for the backup services (y-axis, first two rows for the partial backup service and the last two rows for the full backup service) and the selected variables and principal components for each backup service (x-axis).

Table E4: Summary of the five-fold cross-validations results of the two proposed models (linear model) and only an intercept (simple) model in all the three stages.

Table E5: Results of the multiple linear regression analyses to estimate final WTP for the partial (left) and full (right) backup service, with and without two outliers.

Table E6: Results of the regression models with varying intercepts by the respondents to estimate final WTP for the partial (left) and full (right) backup service, with and without two outliers.

Figure F1: Boxplots of "sure" WTP for the partial (left) and full (right) backup service in the three stages by outage experiences; left of each outage experience group: initial WTP; middle of each outage experience group: middle WTP; right of each outage experience group: final WTP).

Figure F2: Value per kWh of high (left) and lower priority loads (right), by outage experiences.

Figure F3: WTP against longest outage experience from three outage experience groups.

Table F1: Summary of the partial and full backup WTP, by outage experiences.

Table F2: Summary of KS tests to compare three groups' WTP for the partial and full backup service.

Table F3: Summary of Wilcoxon signed-rank tests to track WTP changes as the survey progressed, within each outage experience group.

Table F4: Summary of the value per kWh of high and lower priority loads, by outage experiences.

Table F5: Summary of KS tests to compare three outage experience groups' value per kWh.

Table F6: Summary of Wilcoxon signed-rank tests to track value per kWh changes as the survey progressed, within each outage experience group.

Table G1: Summary of the "sure" value per kWh of high and lower priority loads using two different thresholds.

Table G2: Summary of the paired Wilcoxon signed-rank tests comparing "sure" value per kWh of high priority and lower priority loads using two different thresholds.