Journal of Food Distribution Research
Volume 57, Issue 1, pp. 9–24

Are Americans Willing to Pay to Improve Food Access for Others? Evidence from a Real Donation Experience

Xuesen Fan[a], Jerrod Penn[b], and Wuyang Hu[c]①

[a]*PhD Student, Department of Agricultural, Environmental, and Developmental Economics, 2120 Fyffe Road, The Ohio State University, Columbus, OH 43210, USA*

[b]*Martin D. Woodin Endowed Associate Professor, Department of Agricultural Economics and Agribusiness, 221 Martin D. Woodin Hall, Louisiana State University, Baton Rouge, LA 70803, USA*

[c]*Professor, Department of Agricultural, Environmental, and Developmental Economics, 2120 Fyffe Road, The Ohio State University, Columbus, OH 43210, USA*

Abstract

We quantify Americans' willingness to donate to Feeding America by comparing hypothetical and real-payment decisions within an identical choice context. Respondents completed six discrete choice sets varying donation amounts ($1, $3, $5) and match-based impact. Mixed logit estimates reveal cost sensitivity and positive valuation of charitable impact in both groups. Hypothetical elicitation overstates giving: donate probability is 0.580 versus 0.458 ($1.55 versus $1.20). This 30% hypothetical bias equals $0.50 per respondent, about five meals. Formal tests confirm no significant differences in preference parameters or scale. These findings provide a real-payment benchmark and 1.30× calibration factor for stated-preference measures in food-access contexts.

Keywords: charitable giving, discrete choice experiment, food insecurity, hypothetical bias, incentive compatibility, scale equivalence, stated preferences

①Corresponding author: Email: Hu.1851@osu.edu

Introduction

Food insecurity affects approximately 44 million Americans annually, including 13 million children (Rabbitt et al., 2023). Feeding America, the nation's largest hunger-relief organization, operates roughly 200 food banks and 60,000 pantries, with individual donations representing a major source of its revenue (Feeding America, 2022). For organizations that depend on donations, two questions are critical: "How much are Americans truly willing to donate?" and "Can survey-based measures be trusted?"

Measuring true donation willingness through surveys presents challenges. The common approach is to ask about donation intentions through a questionnaire, but when choices carry no real monetary consequences, stated intentions tend to exceed actual behavior—a phenomenon termed hypothetical bias (Diamond and Hausman, 1994; Penn, Hu, and Ye, 2024). A meta-analysis by Penn and Hu (2018) finds that hypothetical willingness-to-pay (WTP) overstates real payments by a factor of 2.29 on average. In the charitable giving domain, List and Gallet (2001) find overstatement of approximately 35%, though evidence remains limited. Potential mechanisms for hypothetical bias include: the psychological satisfaction from merely expressing donation intentions (Andreoni, 1990), social desirability bias (Krosnick, 1999), or the more broadly defined insufficient consideration of budget constraints in hypothetical elicitation (Penn and Hu, 2018). Accurately measuring donation willingness usually requires real-payment experiments.

However, existing real-payment research in charitable giving has limitations. First, evidence on hypothetical bias comes primarily from environmental goods or international disaster relief (e.g., Champ and Bishop, 2001; Penn, Hu, and Ye, 2024), leaving high-salience domestic causes like hunger relief underexplored. Second, many studies vary both incentive structure and choice set context simultaneously, confounding the pure effect of hypothetical versus real-payment elicitation (e.g., Fang et al., 2021; Lusk and Schroeder, 2004). Third, few studies provide domain-specific calibration factors applicable to similar settings.

This study addresses these gaps. We use a discrete choice experiment (DCE) to elicit donation preferences for Feeding America under both hypothetical and real-payment elicitation, employing a between-subjects design with identical choice context. The experiment varies two attributes: donation amount (\$1, \$3, \$5) and match-generated additional impact—chosen based on findings from the charitable giving literature that donations are cost-sensitive (Clotfelter, 1985) and that donors positively value match-amplified impact (Eckel and Grossman, 2003; Karlan and List, 2007). Our contributions are threefold: (i) we provide incentive-compatible estimates of willingness to donate to hunger relief; (ii) we offer clean identification of hypothetical bias by holding choice context constant; and (iii) we quantify the bias magnitude and provide a calibration factor. In addition, because we implemented our study online without the need for researchers to interact with the respondents in the process, we are able to expand the literature by performing fully automated online real experiments. Results show that under real-payment conditions, donate probability is 0.458 with average donation of \$1.198; the hypothetical treatment group overstates by approximately 30%, yielding a 1.30× calibration factor. Formal scale and preference tests further confirm that the two groups are statistically indistinguishable in their underlying parameters.

Methodology

Survey Design and Administration

Our study uses a DCE survey conducted in May 2021 through Qualtrics, an online survey platform. The survey elicits consumer preferences for donations to Feeding America. Following Carson and Groves (2007), our DCE uses repeated binary choices, presenting a single alternative plus an opt-out option per choice set. Evidence indicates that repeated binary choices have the advantage of being incentive-compatible, enhancing the probability of eliciting truthful preferences (Carson, Groves, and List, 2014; Vossler, Doyon, and Rondeau, 2012;).

In each choice set, the respondent decides whether or not to donate based on the donation profile shown, which is composed of three attributes: (i) the amount to be donated (*Contribute*), (ii) the amount of a match to receive from the researcher and decided by a multiplier (*Multiplier*), and (iii) the chance of receiving that match (*Chance*). The survey randomly displays one of the three contribution amounts: \$1, \$3, or \$5. If the respondents choose to contribute, they will be entered into a lottery to win the match. The lottery has two features: the amount of match—additional funds donated to the charity by the researcher through the multiplier (0, 0.5×, 1×, 2×)—and the chance of receiving that match (0%, 20%, 40%, 60%). If the lottery is won, the respondent's contribution will be matched by the researchers with the multiplier, generating a total donation equal to *Contribute* + *Contribute* × *Multiplier*. If the lottery is not won, the donation amount *Contribute* will not change.

Figure 1 illustrates an example. If respondents choose to contribute, the \$3 contribution amount will be entered into a draw with a 20% chance of winning. If they win the drawing, the \$3 contribution is matched by the winning multiplier (\$3 × 2), generating an extra \$6 donation, with a \$9 donation in total to Feeding America; otherwise, the total donation is \$3. Each choice set incorporates varying attribute levels, and Table 1 details the attributes and their corresponding levels. The design of these attributes is informed by a related study examining risk aversion behaviors (Penn, Howard, and Hu, 2024).

If you choose to contribute, imagine that the contribution amount \$k will be subtracted from your \$5. Would you contribute to the amount below to Feeding America?

Donation Choice Task

Contribution Amount: \$3

Winning Multiplier: 2 x (extra \$6 donated)

Chance of Winning: 20%

- Contribute
- Do not contribute

Figure 1. A Sample Choice Set

Table 1. Discrete Choice Experiment Attributes and Levels

Attribute	Levels	Description
Contribute	\$1, \$3, \$5	Amount of money to support the targeted organization
Multiplier	0, 0.5, 1, 2	The multiplier of additional donations if the respondent wins the lottery
Chance	0%, 20%, 40%, 60%	The probability to win the lottery

Our design also allows explicit capture of risk preferences. However, while one can follow relevant work to incorporate various forms of risk preferences (Harrison et al., 2014; Wu, Mentzakis, and Schaafsma, 2022), they are not the focus of this study. We therefore follow the standard assumption of risk neutrality to isolate the effect of the matching mechanism. We construct a variable, *Match*, defined as the product of *Contribute*, *Multiplier*, and *Chance*, representing the expected dollar value of the additional donation in the lottery. This variable reflects the anticipated researcher-provided match amount based on the probability of winning. *Match* is mathematically equivalent to including a *Multiplier* and *Chance* separately. We also conduct robustness checks using *Match* = *Multiplier* × *Chance* and *Multiplier* and *Chance* separately.

To create the choice sets, we apply a standard blocked full-factorial design (Louviere et al., 1999). A total of six blocks of six repeated binary choice sets was generated. The DCE design contained a restriction such that *Multiplier* = 0 and *Chance* = 0 must always occur together.

Between-Subject Treatment Design

To test hypothetical bias, we need to implement both a hypothetical and a real-payment elicitation. We adopt a between-subject design, where respondents were randomly assigned to either a hypothetical or a real-payment treatment group. This method helps avoid potential ordering and anchoring effects from a within-respondent sequential design (Day et al., 2012; Penn and Hu, 2021). Respondents in both the hypothetical treatment group and real-payment treatment group received a \$5 participation payment. The initial sample included 319 (hypothetical) and 312 (real-payment) respondents. After removing respondents who failed the attention-check question introduced by Malone and Lusk (2019) (see Appendix for the attention-check question), the final sample consists of 267 respondents (hypothetical) and 277 respondents (real-payment). Table 2 reports sample demographics by treatment group; the two groups are broadly comparable on observables. In the hypothetical treatment group, choices are hypothetical: each respondent completes a DCE comprising one of the six blocks of choice sets randomly assigned to them. Each block contains six binary choice sets pertaining to hypothetical elicitation, in which the DCE informs the respondent prior to the choice tasks that no actual payment occurs regardless of the choice made.

Table 2. Sample Characteristics

Variable	Hypothetical Sample (N = 267)	Real Sample (N = 277)	Difference (Hypo-Real)	2020 ACS (Adults 18+)
Age (%)				
18–24	10.76	8.40	2.36	12.56
25–34	22.15	21.37	0.78	17.94
35–44	18.99	18.58	0.41	16.26
45–54	13.29	13.49	-0.20	16.39
55–64	15.19	14.25	0.94	16.65
65–74	13.29	17.30	-4.01	12.07
75–84	5.70	5.85	-0.15	6.06
85+	0.63	0.76	-0.13	2.58
Female (%)	48.73	52.93	-4.20	50.76
Race (%)				
American Indian/Alaskan Native	1.23	1.72	-0.49	0.68
Asian	2.47	3.93	-1.46	7.70
Black or African American	14.20	12.29	1.91	12.05
Native Hawaiian/Pacific Islander	0.00	0.25	-0.25	0.19
White	79.63	78.62	1.01	76.53
Other	2.47	2.46	0.01	0.51
Prefer not to answer	0.00	0.25	-0.25	—
Education (%)				
Some high school or less	0.65	0.77	-0.12	11.55
High school diploma/GED	15.58	14.69	0.89	27.32
Some college/2-year degree/technical school	35.06	35.82	-0.76	30.74
Bachelor's degree	33.12	33.25	-0.13	19.20
Graduate/advanced degree	14.94	15.21	-0.27	11.18
Prefer not to answer	0.65	0.26	0.39	—
Income (%)				
Under $15,000	9.49	7.12	2.37	9.90
$15,000 to $24,999	8.23	8.40	-0.17	8.50
$25,000 to $34,999	8.86	11.45	-2.59	8.60
$35,000 to $49,999	11.39	12.98	-1.59	12.00
$50,000 to $74,999	20.89	21.12	-0.23	17.20
$75,000 to $99,999	15.19	14.76	0.43	12.80
$100,000 to $149,999	17.09	14.25	2.84	15.60
$150,000 to $199,999	3.80	3.82	-0.02	7.10
$200,000 or more	3.16	3.31	-0.15	8.30
Prefer not to answer	1.90	2.80	-0.90	—

Notes: Entries are percentages. Difference (Hypo − Real) reports values in the hypothetical treatment group minus those in the real-payment treatment group. Population shares are from the 2020 American Community Survey (ACS) for adults aged 18 and above. "—" denotes not collected or not available.

In the real-payment treatment group, each respondent is randomly assigned to one of the same six blocks of choice sets, except that among the six choice sets, one of the sets will be randomly chosen by the survey system as binding. Respondents will make an actual donation according to their decision recorded in the binding choice set. If a respondent chose to donate in the binding choice set, the specified amount was deducted from the $5 payment and transferred to Feeding America. Participants must acknowledge and agree to a statement confirming they understand that one and only one of the six choice sets will be randomly selected as binding and their decision recorded in that choice set will be implemented.

Empirical Model

Our empirical strategy is grounded in the random utility model. We employ a mixed logit model to analyze respondent choices, allowing for the incorporation of unobserved preference heterogeneity across individuals. Each respondent completes repeated binary donation choices. Normalizing the utility associated with the no-donate option to zero, utility from donating in choice set *t* for respondent *n* is:

$$U_{nt} = \alpha + \beta_1 Match_{nt} + \beta_2 \ln(1 + Contribute_{nt}) + \epsilon_{nt} \tag{1}$$

with ϵ_{nt} i.i.d. Type I largest extreme value. The log specification captures diminishing marginal disutility of cost. Following the suggestion of Hu et al. (2022), we dummy-code *Donate* alternative-specific constant (ASC) α. We then treat it as a fixed parameter across the respondents and specify (β_1, β_2) as random coefficients. The model is estimated by simulated maximum likelihood with standard errors clustered at the respondent level.

With the log cost specification, marginal willingness-to-pay (MWTP) for the *Donate* ASC and *Match*, evaluated at contribution level *c*, is：

$$MWTP_{Donate}(c) = -\frac{\alpha}{\beta_2}(1 + c) \tag{2}$$

$$MWTP_{Match}(c) = -\frac{\beta_1}{\beta_2}(1 + c) \tag{3}$$

We report MWTP at *Contribution* = $3 and assess robustness at *Contribution* = $1 and $5. MWTP standard errors are obtained via the delta method using the estimated variance–covariance matrix.

To connect preference estimates to donation behavior, we compute model-implied donate probabilities and expected donations at the choice sets level and report sample means by treatment group.

Results

Preference Estimates

Table 3 reports mixed logit estimates. The *Donate* ASC is positive and significant in both treatment groups, indicating general respondent support to donating to Feeding America. The coefficient on *ln(1+Contribute)* is negative and significant, confirming cost sensitivity. *Match* enters positively and significantly, implying respondents value match-generated impact. Because models under the hypothetical treatment group and real-payment treatment group are estimated separately, coefficient magnitudes should not be compared directly due to potential scale differences. We therefore base comparisons on MWTP. To formally test whether preference parameters differ across treatment conditions, we estimate a pooled mixed logit model with treatment interaction terms. None of the interaction terms are statistically significant (*Hypo* × *Donate*: p = 0.501; *Hypo* × *Match*: p = 0.408; *Hypo* × ln(1+*Contribute*): p = 0.749), indicating no significant preference differences between the hypothetical and real-payment groups. We further implement the Swait and Louviere (1993) scale-adjusted test, which confirms that neither scale (μ^* = 0.80, p = 0.443) nor preference parameters (p = 0.753) differ significantly between the two groups. These results validate our MWTP-based comparison, as MWTP derived from coefficient ratios is invariant to multiplicative scale differences. Nonetheless, the absence of significant parameter differences does not imply identical MWTPs: because MWTP is a nonlinear ratio of parameters, opposite-direction deviations across the numerator and denominator compound into the observed 30% behavioral gap. Additionally, a Poe test confirms that neither the *Donate* ASC MWTP (p = 0.251) nor the *Match* MWTP (p = 0.299) differs significantly between the two groups, further supporting the conclusion of minimal hypothetical bias at the preference level.

Table 3. Mixed Logit Estimates for Donation Choices: Hypothetical vs. Real

Variables	**Hypothetical**	**Real**
Panel A. mean coefficients		
Donate	1.638**	1.378***
	(0.692)	(0.480)
Match	0.662*	0.481**
	(0.404)	(0.203)
ln(1 + contribute)	-1.553**	-1.599***
	(0.748)	(0.513)
Panel B. standard deviations of random coefficients		
SD(match)	0.705	0.045
	(1.194)	(0.189)
SD[ln(1 + contribute)]	2.296***	2.884***
	(0.789)	(0.499)
# of respondents	267	277
# of observations	1,602	1,662
Log likelihood	-198.163	-173.356
McFadden pseudo R2	0.0984	0.0724

Notes: Standard errors (in parentheses) are cluster-robust at the respondent level. The base alternative is No-donate (ASC corresponds to *Donate*). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. Observations refer to individual choice sets.

Willingness-to-Pay

Table 4 reports MWTP evaluated at *Contribute* = $3. MWTP for the *Donate* ASC is $4.22 in the hypothetical treatment group versus $3.45 in the real-payment treatment group. MWTP for *Match* is $1.71 versus $1.20 in the two respective treatment groups. Both donation propensity and valuation of match impact are higher under hypothetical elicitation, consistent with upward hypothetical bias predicted by the literature.

Table 4. Marginal Willingness to Pay Estimates: Hypothetical vs. Real

Attribute	Hypothetical	Real	Poe Test (*p*-value)
Donate	4.217***	3.447***	0.251
	(0.911)	(0.657)	
Match	1.705**	1.201***	0.299
	(0.698)	(0.436)	

Notes: Entries are marginal WTP in USD, implied by Table 3 and evaluated at *Contribute* = $3. Standard errors (reported in parentheses) are obtained using the delta method based on the estimated variance–covariance matrix. Hypothetical and Real denote hypothetical and real-payment elicitation, respectively. The last column reports one-sided *p*-values from the Poe, Giraud, and Loomis (2005) test of the null hypothesis that hypothetical MWTP equals real MWTP (Hypo/Real = 1); a *p*-value greater than 0.10 indicates failure to reject equality. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Predicted Donation Behavior

Table 5 translates estimates into behavioral predictions. We report three outcome measures: donate probability (extensive margin), expected out-of-pocket donation (intensive margin), and expected total dollars reaching Feeding America (actual charitable impact). Mean donate probability is 0.580 in the hypothetical treatment group versus 0.458 in the real-payment treatment group, a 27% overstatement. Expected out-of-pocket donations are $1.55 versus $1.20 in the two respective treatment groups, a 30% overstatement. Expected total dollars are $2.23 versus $1.73 in the hypothetical treatment group and real-payment treatment group, respectively—a gap of $0.50 per respondent, or approximately 5 meals using Feeding America's conversion ($1 ≈ 10 meals). These results establish that: (i) Americans exhibit substantial real willingness to donate to hunger relief (45.8% probability; and about $1.20 average); (ii) hypothetical elicitation overstates donation behavior by approximately 30%; and (iii) the implied calibration factor is 1.30×.

Table 5. Model-Implied Donation Probabilities and Expected Donations: Hypothetical vs. Real

Outcome	Hypothetical	Real	Difference
Mean donate probability	0.580	0.458	0.122
Expected out-of-pocket donation (USD)	1.554	1.198	0.356
Expected total donation (USD)	2.228	1.727	0.501
Expected total donation (meals)	22.284	17.274	5.010

Notes: Entries are model-implied sample means computed separately for hypothetical treatment group and real-payment treatment group. “Expected out-of-pocket donation” is the predicted amount paid by the respondent. “Expected total donation” is the predicted amount received by Feeding America, equal to the respondent’s contribution plus the expected researcher-funded match-lottery add-on (“*Match*”). Difference equals Hypothetical minus Real. $1 = 10 meals.

Robustness

Appendix tables confirm robustness. Table A1 shows MWTP patterns held at alternative contribution levels ($1 and $5), with the hypothetical treatment group consistently higher. Tables A2–A3 replace *Match* = *Multiplier* × *Chance* × *Contribution* with *Match* = *Multiplier* × *Chance*; results are qualitatively unchanged. Tables A4–A5 enter *Multiplier* and *Chance* separately; estimates are less precise due to increased parameterization, supporting our use of *Match* as a more efficient specification. Across all specifications, the hypothetical-real gap persists.

To examine whether hypothetical bias varies across sociodemographic subgroups, we estimate a pooled mixed logit model with three-way interaction terms (*Donate* × *Hypo* × *Demographic*). The results (see Table A6) show that hypothetical bias in donation propensity is significantly larger among low-income respondents (coefficient = 1.749, $p = 0.040$), whereas no significant differences are found for gender ($p = 0.266$) or education ($p = 0.506$). This finding suggests that the overall hypothetical bias is driven disproportionately by low-income respondents, who face stronger budget constraints under real-payment conditions. These findings diverge from (Mjelde et al., 2012), who found that education—but not income or gender—significantly influences the hypothetical bias ratio across three South Korean contingent valuation studies. While both studies find no gender effect, income is significant in our context of charitable giving, whereas education does so in theirs, suggesting that demographic determinants of hypothetical bias may be context-dependent.

Conclusion

This study quantifies Americans' real willingness to donate to hunger relief and the magnitude of hypothetical bias. Using an online discrete choice experiment with hypothetical and real-payment elicitation implemented through a random-binding mechanism, we find two main results. First, respondents exhibit clear willingness to donate: the real-payment elicitation donate probability is 0.458, indicating substantial baseline support for Feeding America. Second, hypothetical choices overstate giving in that model-implied participation in donation, out-of-pocket donations, and estimated total dollars reaching the charity are all higher in the hypothetical elicitation. The 30% overstatement corresponds to approximately five meals per respondent. By holding choice sets constant across treatment groups, we provide clean identification of hypothetical bias and a 1.30× calibration factor. Notably, this calibration factor is relatively modest compared to the meta-analytic average of 2.29× reported by Penn and Hu (2018). One possible explanation for this finding is respondents' relatively high familiarity with Feeding America as a nationally recognized hunger-relief organization; greater familiarity with a charitable cause may stabilize preferences and reduce hypothetical bias relative to less familiar goods.

These findings address the three gaps identified in the literature. First, existing evidence on hypothetical bias comes primarily from environmental goods or international disaster relief; we extend this research to hunger relief, a high-salience domestic cause. Second, many studies confound incentive structure and choice set context; by holding choice set context constant between real-payment and hypothetical elicitation, we isolate the effect of hypothetical bias. Third, we provide a domain-specific calibration factor directly applicable to food-access contexts. Additionally, our fully automated online implementation demonstrates that real-payment experiments can be conducted at scale without researcher-respondent interaction.

For practitioners, our findings suggest that survey-based donation amounts to be deflated by approximately 30% (or divided by 1.30) when projecting actual giving. For charitable organizations like Feeding America, this calibration can improve operational planning and resource allocation. For researchers, the results reinforce the importance of incentive-compatible elicitation when measuring charitable preferences with policy or operational implications.

This study has limitations. Our real-payment elicitation is capped at $5, which may dampen observed cost sensitivity relative to uncapped settings. Consequently, the 1.30× calibration factor is estimated within this low-stakes context, and its applicability to higher-donation settings remains an open empirical question. While our between-subjects design eliminates order effects, it precludes within-person comparisons. Such analysis can be important for understanding how individual preferences shift between hypothetical and real decisions. Future research could also implement higher budgets and validate predictions from hypothetical elicitations against field fundraising outcomes. Additionally, future work could examine whether hypothetical bias varies across charitable causes or donor populations. Our analysis of demographic heterogeneity reveals that hypothetical bias is significantly larger among low-income respondents, while no significant differences are found for gender or education. The precision of this type of subgroup-specific calibration factors would benefit from larger samples designed with stratified recruitment across demographic cells.

Data Availability Statement

The data used in this study were collected from a private survey and contain confidential respondent information. Due to privacy and confidentiality agreements, the data are not publicly available but can be provided by the corresponding author upon reasonable request.

References

Andreoni, J. 1990. “Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving.” *The Economic Journal* 100(401):464–477.

Carson, R., and T. Groves. 2007. “Incentive and Informational Properties of Preference Questions.” *Environmental and Resource Economics* 37:181–210.

Carson, R.T., T. Groves, and J.A. List. 2014. “Consequentiality: A Theoretical and Experimental Exploration of a Single Binary Choice.” *Journal of the Association of Environmental and Resource Economists* 1(1/2):171–207.

Champ, P.A., and R.C. 2001. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias." *Environmental and Resource Economics* 19(4):383–402.

Clotfelter, C.T. 1985. "Charitable Giving and Tax Legislation in the Reagan Era." *Law and Contemporary Problems* 48(4):197–212.

Day, B., I.J. Bateman, R.T. Carson, D. Dupont, J.J. Louviere, S. Morimoto, R. Scarpa, and P. Wang. 2012. "Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies." *Journal of Environmental Economics and Management* 63(1):73–91.

Diamond, P.A., and J.A. Hausman. 1994. "Contingent Valuation: Is Some Number Better Than No Number?" *Journal of Economic Perspectives* 8(4):45–64.

Eckel, C.C., and P.J. Grossman. 2003. "Rebate versus Matching: Does How We Subsidize Charitable Contributions Matter?" *Journal of Public Economics* 87(3–4):681–701.

Fang, D., R.M. Nayga, G.H. West, C. Bazzani, W. Yang, B.C. Lok, C.E. Levy, and H.A. Snell. 2021. "On the Use of Virtual Reality in Mitigating Hypothetical Bias in Choice Experiments." *American Journal of Agricultural Economics* 103(1):142–161.

Feeding America. 2022. "A Bold Aspiration—2022 Annual Report Impact Report." *Feeding America*. Available online: https://www.feedingamerica.org/about-us/financials.

Harrison, M., D. Rigby, C. Vass, T. Flynn, J. Louviere, and K. Payne. 2014. "Risk As an Attribute in Discrete Choice Experiments: A Systematic Review of the Literature." *The Patient-Patient-Centered Outcomes Research* 7:151–170.

Hu, W., S. Sun, J. Penn, and P. Qing. 2022. "Dummy and Effects Coding Variables in Discrete Choice Analysis." *American Journal of Agricultural Economics* 104(5):1770–1788.

Karlan, D., and J.A. List. 2007. "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment." *American Economic Review* 97(5):1774–1793.

Krosnick, J.A. 1999. "Survey Research." *Annual Review of Psychology* 50(1):537–567.

List, J.A., and C.A. Gallet. 2001. "What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20:241–254.

Louviere, J.J., R.J. Meyer, D.S. Bunch, R. Carson, B. Dellaert, W.M. Hanemann, D. Hensher, and J. Irwin. 1999. "Combining Sources of Preference Data for Modeling Complex Decision Processes." *Marketing Letters* 10:205–217.

Lusk, J.L., and T.C. Schroeder. 2004. "Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks." *American Journal of Agricultural Economics* 86(2):467–482.

Malone, T., and J.L. Lusk. 2019. "Releasing the Trap: A Method to Reduce Inattention Bias in Survey Data with Application to US Beer Taxes." *Economic Inquiry* 57(1):584–599.

Mjelde, J.W., Y.H. Jin, C.-K. Lee, T.-K. Kim, and S.-Y. Han. 2012. "Development of a Bias Ratio to Examine Factors Influencing Hypothetical Bias." *Journal of Environmental Management* 95(1):39–48.

Penn, J., G. Howard, and W. Hu. 2024. "Model Choice, Hypothetical Bias and Risk Aversion: A Charitable Donation Application." Working paper.

Penn, J., W. Hu, and T. Ye. 2024. "Efficacy of Hypothetical Bias Mitigation Techniques: A Cross-Country Comparison." *Journal of Environmental Economics and Management* 125: 102989.

Penn, J.M., and W. Hu. 2018. "Understanding Hypothetical Bias: An Enhanced Meta-Analysis." *American Journal of Agricultural Economics* 100(4):1186–1206.

Penn, J.M., and W. Hu. 2021. "The Extent of Hypothetical Bias in Willingness to Accept." *American Journal of Agricultural Economics* 103(1):126–141.

Poe, G.L., K.L. Giraud, and J.B. Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87(2):353–365.

Rabbitt, M.P., L.J. Hales, M.P. Burke, and A. Coleman-Jensen. 2023. "Household Food Security in the United States in 2022." Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Swait, J., and J. Louviere. 1993. "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models." *Journal of Marketing Research* 30(3):305–314.

Vossler, C.A., M. Doyon, and D. Rondeau. 2012. "Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments." *American Economic Journal: Microeconomics* 4(4):145–171.

Wu, H., E. Mentzakis, and M. Schaafsma. 2022. "Exploring Different Assumptions about Outcome-Related Risk Perceptions in Discrete Choice Experiments." *Environmental and Resource Economics* 81(3):531–572.

Appendix: Attention Check Question

Recent research on decision making shows that choices are affected by context. Differences in how people feel, their previous knowledge and experience, and their environment can affect choices. To help us understand how people make decisions, we are interested in information about you. Specifically, we are interested in whether you actually take the time to read the directions; if not, some results may not tell us very much about decision making in the real world. To show that you have read the instructions, please ignore the question below about how you are feeling and instead check only the 'none of the above' option as your answer. Please check all the words that describe how you are currently feeling."

- o Fear
- o Anger
- o Sadness
- o Joy
- o Disgust
- o Surprise
- o Trust
- o Hope
- o None of the above

Table A1. MWTP Robustness by Donation Amount (Contribute = $1, $3, $5): Hypothetical vs. Real

	Donate (USD)		Match (USD)	
Contribute ($)	**Hypothetical**	**Real**	**Hypothetical**	**Real**
1	2.108*** (0.456)	1.724*** (0.326)	0.853** (0.349)	0.601*** (0.218)
3	4.217*** (0.911)	3.447*** (0.652)	1.705** (0.698)	1.201*** (0.436)
5	6.325*** (1.367)	5.171*** (0.977)	2.558** (1.048)	1.802*** (0.654)

Notes: Entries are MWTP (USD). Standard errors in parentheses are computed using the delta method from the estimated variance–covariance matrix. Hypothetical and Real correspond to hypothetical treatment group and real-payment treatment group. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A2. Robustness Check: Mixed Logit Estimates (Match = Multiplier × Chance): Hypothetical vs. Real

Variables	Hypothetical	Real
Panel A. Mean coefficients		
Donate	1.487**	0.905*
	(0.697)	(1.226)
Match	0.0251*	0.0152**
	(0.0138)	(0.00598)
ln(1 + contribute)	-1.409**	-1.255***
	(0.643)	(0.457)
Panel B. standard deviations of random coefficients		
SD(match)	0.0842***	2.941***
	(0.031)	(0.513)
SD[ln(1 + contribute)]	1.736**	-0.00194
	(0.69)	(0.0154)
# of respondents	267	277
# of observations	1,602	1,662
Log likelihood	-200.174	-189.365
McFadden pseudo R2	0.143	0.0762

Notes: Standard errors (in parentheses) are cluster-robust at the respondent level. The base alternative is No-donate (ASC corresponds to *Donate*). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A3. Robustness Check: MWTP Estimates (Match = Multiplier × Chance): Hypothetical vs. Real

Attribute	Hypothetical	Real
Donate	4.223***	2.886***
	(0.991)	(0.930)
Match	7.123**	4.845**
	(3.554)	(2.149)

Notes: Entries are marginal WTP in USD implied by Table A1 and evaluated at *Contribute* = $3. Standard errors (reported in parentheses) are obtained using the delta method based on the estimated variance–covariance matrix. Hypothetical and Real denote hypothetical treatment group and real-payment treatment group, respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table A4. Robustness Check: Mixed Logit Estimates (Multiplier and Chance): Hypothetical vs. Real

Variables	Hypothetical	Real
Panel A. mean coefficients		
Donate	2.547	0.779
	(1.681)	(0.603)
Multiplier	0.842	0.519*
	(0.727)	(0.301)
Chance	0.00639	0.0132
	(0.0179)	(0.00870)
ln(1 + contribute)	-2.297	-1.489**
	(1.674)	(0.584)
Panel B. standard deviations of random coefficients		
SD(multiplier)	4.156*	0.995
	(2.498)	(1.388)
SD(chance)	0.0501	2.17e-05
	(0.0529)	(0.00281)
SD[ln(1 + contribute)]	2.264	3.014***
	(1.508)	(0.572)
# of respondents	267	277
# of observations	1,602	1,662
Log likelihood	-197.287	-173.356
McFadden pseudo R2	0.0853	0.0836

Notes: Cluster-robust standard errors (respondent level) in parentheses. The base alternative is No-donate; the ASC corresponds to *Donate*. Standard deviations are reported in absolute value. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A5. Robustness Check: MWTP Estimates (Multiplier and Chance): Hypothetical vs. Real

Attribute	Hypothetical	Real
Donate	4.434***	2.092*
	(0.925)	(1.128)
Multiplier	1.466	1.394**
	(1.0433)	(0.670)
Chance	0.0111	0.0354
	(0.0281)	(0.0277)

Notes: Entries are MWTP (USD) implied by Table A3 and evaluated at *Contribute* = $3. MWTP standard errors (in parentheses) are obtained using the delta method. *Multiplier* is measured in multiplier units; *Chance* is measured in percentage points, so MWTP for *Chance* corresponds to a 1-percentage-point increase in winning probability. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A6. Pooled Mixed Logit with Demographic Heterogeneity in Hypothetical Bias

Variable	Pooled
Panel A. mean coefficients	
Donate	2.257***
	(0.653)
Donate × hypo	-0.988
	(0.742)
Donate × male	-0.583
	(0.492)
Donate × low-income	-1.211**
	(0.508)
Donate × low-edu	-0.556
	(0.509)
Donate × hypo × male	0.879
	(0.789)
Donate × hypo × low-income	1.749**
	(0.850)
Donate × hypo × low-edu	0.539
	(0.811)
Match	0.564***
	(0.214)
ln(1 + contribute)	-1.437***
	(0.456)
Panel B. standard deviations of random coefficients	
SD(Match)	0.316
	(1.026)
SD[ln(1 + Contribute)]	2.707***
	(0.485)
# of respondents	492
# of observations	2,952
Log likelihood	-514.238
McFadden pseudo R2	0.0518

Notes: Standard errors (in parentheses) are cluster-robust at the respondent level. The base alternative is No-donate (ASC corresponds to *Donate*). *Hypo* is a binary indicator equal to one for respondents in the hypothetical treatment group. *Low-income* is defined as household income below $50,000. *Low-edu* is defined as not having a 4-year college degree. *Male* is a binary indicator for male respondents. After excluding respondents who preferred not to report demographics, the estimation sample consists of 492 respondents (2,952 observations). *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.