Web Appendix

for

Average Value Affects Consumer Budgets

Table of Contents

WEB APPENDIX A: SURVEY OF BUDGETING EXPERIENCE

A representative survey by Zhang et al. (2022) suggests consumer budgeting is prevalent. Approximately two out of every three Americans currently use budgets; of those who do not currently budget, 42% have budgeted in the past. Of those who currently budget, 59% use formal budgets. Budgeting is common across income and wealth levels. Consumers typically organize budgets according to categories of spending: The most common labels consumers spontaneously report for their budgets include necessities like rent, mortgage, and insurance, as well as discretionary purchases like dining and entertainment.

We motivate the consumer relevance of our work with a survey of consumers' own budgeting experiences, drawing from and building on Zhang et al. (2022). Using both open-ended and closed-ended survey items, we assess the motivations and strategies for setting, tracking, and following budgets. In particular, we consider how consumers budget for discretionary spending categories. Whereas budgets for necessities are often fixed at specific payment amounts (e.g., rent, recurring bills, debt repayment; as measured in the manuscript study 1), budgets for discretionary purchases are more likely to be set based on consumer preferences, in which case the value of budget categories may play an important role.

Method

We surveyed 200 participants from a gender-balanced sample on Amazon Mechanical Turk (AMT). The survey consisted of 4 open-ended questions and 11 closed-ended budgeting questions, many of which included follow-up components. We adapted the basic structure of

Zhang et al. (2022), which began by establishing the participant's personal experience with budgeting. Following their approach, we dropped from our analysis all observations from participants who indicated no current or prior budgeting experience. Though not analyzed, these participants progressed through the survey by imagining the budget they would keep if they were to start budgeting. The complete survey materials including summary statistics (for the analyzed group of participants with current or prior budgeting experience) are available in our ResearchBox.

Key Findings

Budgets are relevant. The first measure (adopted from Zhang et al. 2022) identifies personal budgeting experience. Overwhelmingly, consumers report using budgets to guide their finances. 72% report currently budgeting, 14% report having previously budgeted, and only 14% report never budgeting (Q1). These percentages are comparable to those from the nationally representative sample used by Zhang et al. (2022), who observe rates of 66%, 15%, and 20% for current budgeters, previous budgeters, and never-budgeters, respectively. In our data, those who have budgeted and those who have never budgeted do not differ in gender, age, educational attainment, or income bracket (ps > .27). Following the approach of Zhang et al., we consider only the responses of the 86% of participants who currently or previously budgeted. All subsequent figures use this 86% of respondents with budgeting experience as the denominator, unless specified otherwise.

The widespread use of budgets in our sample reflects a variety of different financial motivations and goals. When asked why they budget (Q5), some participants used budgets to overcome challenges of self-control (e.g., "*I like to make sure I don't do anything crazy or develop bad spending habits*"; "*I need to budget or I will end up overspending*.") Others were motivated by simplicity (e.g., "*I don't want to worry about money. I want to set aside money into each pool and then spend whatever I have left and not worry about retirement or debt or anything*.") Some articulated goals for spreading consumption across categories (e.g., "*I budget money so that I know how much money I have and I can allocate it to different needs. I can also save money for specific things instead of just having one large lump sum*"*;* "*I budget my money because I like to do things like go to movies, buy clothes, and go to restaurants; but if I don't budget towards these things, then I'll end up spending way too much on these non-necessities, then not have enough towards my bills.*") Regardless of the motivations and goals for budgeting,the act of budgeting creates a categorical structure for evaluating potential expenditures, which may in turn impact consumption.

Budgets are clearly defined and frequently checked. The majority of budgeters (67%) use some type of formal budgets to record and update transactions, compared to the 33% who rely solely on informal budgets (i.e., mental accounts; Thaler 1980, 1985, 1999) to keep track of finances (Q2a). Among those practicing formal budgeting, the most common approaches were pen-and-paper budgeting (37% of respondents) and computer spreadsheets (33% of respondents), followed by budgeting apps (12%) and websites (5%) (Q2b). Consumers regularly monitor their formal budgets, with 57% of those with formal budgets checking at least every few days and 89% checking at least every week (Q3). This high frequency of checking is not random, but rather reflects consulting budgets prior to spending. In our survey, 98% of consumers prefer to consult their budgets prior to making a purchase, compared to only 2% who prefer to check after making a purchase (Q15). Taken together, these observations suggest budgets are clearly defined, regularly checked, and checking a budget is a precursor to spending. The implication is that budgets will guide spending.

Budgets are consequential for spending across categories. Budgets impact consumption when they are followed. When asked about the importance of following one's budget (1 = "Not very important"; 7 = "Very important"), the modal response was the maximum of 7 (M = 6.18, SD 0.93) (Q14). To explore the importance of distinct budget categories, we modified a question about the main reasons for budgeting (Zhang et al. 2022: question 5). As an additional potential reason, we added: "to make sure I know how much is available to spend in different categories" (Q6). Critically, a majority (58%) of respondents indicated this as one of the main reasons for budgeting. In fact, of the 10 possible reasons, only two had higher response rates (table A.1). We take this as evidence that the multicategorical nature of budgets is an important and appealing aspect of using budgets. In other words: Many consumers are drawn to budgets precisely to guide allocation across different categories. And these allocations are followed. When imagining unexpected budget deviations, 82% of respondents indicated they would rather reduce spending within the overspent category than rebalance their allocations across budgets (Q13). Budgets are sticky and have a direct consequence for how people spend.

TABLE A.1
MAIN REASONS FOR BUDGETING

	Response	Frequency	Proportion
1	To make sure I don't spend more than my income	139	0.81
2	**To make sure I know how much is available to spend in different categories**	99	0.58
3	To save for long-term goals	110	0.64
4	To save for short-term goals	85	0.50
5	To avoid debt from predictable overspending	83	0.49
6	To avoid debt from unforeseen expenses	86	0.50
7	To make sure that I can provide for my family	83	0.49
8	To get myself out of debt	51	0.30
9	Other	7	0.04
10	I don't think it is important to budget	1	0.01

Note—The "main reasons for budgeting," modified from Zhang et al. (2022): question 5. We introduce and are interested in the response in the second line: "to make sure I know how much is available to spend in different categories." The asterisks are added here for visual emphasis but were not included in the stimuli presented to participants. Participants could select multiple responses.

These responses suggest consumers use budgets to guide their spending across different categories. How are budget categories structured? Adopting a question from Zhang et al. (2022), about 10% of consumers prefer extremely coarse categories (i.e., "necessities, discretionary") and 20% prefer extremely granular categories (i.e., "rent, utilities, cell phone, internet, car, groceries, dining out, movies, travel, clothing, exercise, healthcare, other"). The remainder of consumers fall somewhere between these two extremes (Q8). While there is substantial heterogeneity across the various levels of responses, the majority (62%) budget at a level of detail that is sufficiently granular to separate Dining Out vs. Entertainment (rows 4-6 of table A.2).

Additionally, we asked participants to list their own budget categories using an open- ended format (Q7). The most common self-generated category labels include "*food,*" "*rent,*" "*utilities,*" "*groceries,*" "*insurance,*" "*gas,*" "*car,*" and "*entertainment.*" Many of these labels refer to fixed-expense categories. Because the budgets for fixed expenses should have little to no variation in allocation or spending over the short run, we focus on discretionary spending categories. Therefore, in the studies in the main text, we consider "Dining" (a label encompassing discretionary elements of "food" and "groceries") and "Entertainment."

TABLE A.2
GRANULARITY OF BUDGETS

	Response	Frequency	Proportion	Zhang et al.
1	Necessities, Discretionary	17	0.10	0.12
2	Housing & Transportation, Food, Discretionary, Other	25	0.15	0.18
3	Housing & Transportation, Food, Entertainment, Clothing, Other	23	0.13	0.13
4	Housing, Car, Groceries, Dining Out, Entertainment, Clothing, Other	28	0.16	0.16
5	Rent, Utilities, Cell phone, Groceries, Dining Out, Movies, etc...	43	0.25	0.18
6	Rent, Utilities, Cell phone, Internet, Car, Groceries, etc...	35	0.20	0.23

Note—This question adopted from Zhang et al. (2022): question 20. Responses in rows 5 and 6 were truncated in this table for presentation purposes. Participants saw the full granular lists, which are also available in our ResearchBox. We note 62% of respondents use budgets sufficiently granular to distinguish between discretionary categories such as Dining Out and Entertainment. The Zhang et al. proportions are presented in the last column.

Discussion

The survey of budgeting experience explored whether consumers budget, why they do so, and how they set, track, and follow their allocations. The key findings suggest budgets are relevant for most consumers, who formally track and follow their budget allocations. The majority of consumers indicate they use budgets in order to guide their spending across different budget categories. While each consumer uses their own category structure, most have sufficiently granular categories to separate dining and entertainment. Therefore, our paper considers dining and entertainment as two common discretionary budget categories.

WEB APPENDIX B: SUPPLEMENT TO STUDY 1

As this was a descriptive survey of consumers' own behaviors and experiences (and not an experiment with clear *a priori* statistical inferences), we did not preregister this study.

Additional measures: item focus and category focus. The manuscript details the key findings for average value and marginal value and indicates that the results for item focus and category focus are reported here. We decided to concentrate our manuscript discussion on average value and marginal value to enhance clarity to the reader, as we discuss (and test) the role of item-level and category-level evaluations later in the main text. The findings here are compatible with—and certainty not contradictory towards—the preregistered, experimental findings of the supplementary Amazon study (manuscript Appendix).

Design

Following the budgeting exercise using the pie chart, participants self-reported four measures pertaining to their focus on (i) average value, (ii) marginal value, (iii) categories, and (iv) items. For all four measures (average value, marginal value, category focus, item focus), we sampled from four different question variants. As discussed in the manuscript, we took this approach to reduce the likelihood that any conclusions are tied to a unique question wording. The full set of variant wordings is provided in table B.1.

TABLE B.1
VARIANT WORDINGS USED IN STUDY 1

When thinking about setting your budget, to what extent did you find yourself…

Measure	Variant wording	Variant id
Average	…thinking about your overall impression of how much you like each category?	1
	…remembering your general liking of each category?	2
	…comparing your overall enthusiasm for each category?	3
	…relying on your general evaluation of each category?	4
Margin	…thinking about what you could buy with just a little more money (or conversely, what you would lose if you spent a little less)?	1
	…imagining how a small adjustment to one of your budgets could change what you buy?	2
	…weighing the trade-offs between having enough money to buy one thing or the other, but not both?	3
	…considering how giving up one thing might allow you to buy something else?	4
Category	…thinking about categories of expenditures?	1
	…considering trade-offs between categories of purchases?	2
	…imagining a collection of purchases with a shared meaning?	3
	…focusing on the big-picture sense of what each category represents?	4
Item	…thinking about and imagining specific, individual expenses?	1
	…considering particular ways to spend money within a given category?	2
	…paying attention to specific things you could purchase?	3
	…visualizing a concrete thing that you will spend your money on?	4

Question ordering was counterbalanced across two factors. The first factor was whether the measure corresponded to value (average, margin) vs. evaluation mode (category, item). We counterbalanced which set of questions (value vs. evaluation mode) appeared 1st and 2nd, as opposed to 3rd and 4th. The second factor was whether the measure aligned more closely with ensemble perception (average value, category evaluations) or normative principles of decision making (marginal value, item evaluations). Thus, this second factor determined whether the ensemble vs. normative questions were presented first or second, conditional upon the first factor. These two counterbalancing factors were not significant predictors of measure responses, which we assess by regressing each of the 16 measure variants on the contrast-coded counterbalancing variables and their interaction. There are no significant main effects of either counterbalancing factor or the interaction after adjusting for multiple comparison testing (to account for the 48 comparisons: 16 main effects from the first counterbalanced factor, 16 main effects from the second counterbalanced factor, and 16 interactions).

Results

In the manuscript, we report our findings using the complete dataset of 100 participants. As footnoted, there were originally 101 complete observations; however, two observations were tied to the same participant identifier. To preserve independent and naïve responses, we eliminated the second observation linked to this participant.

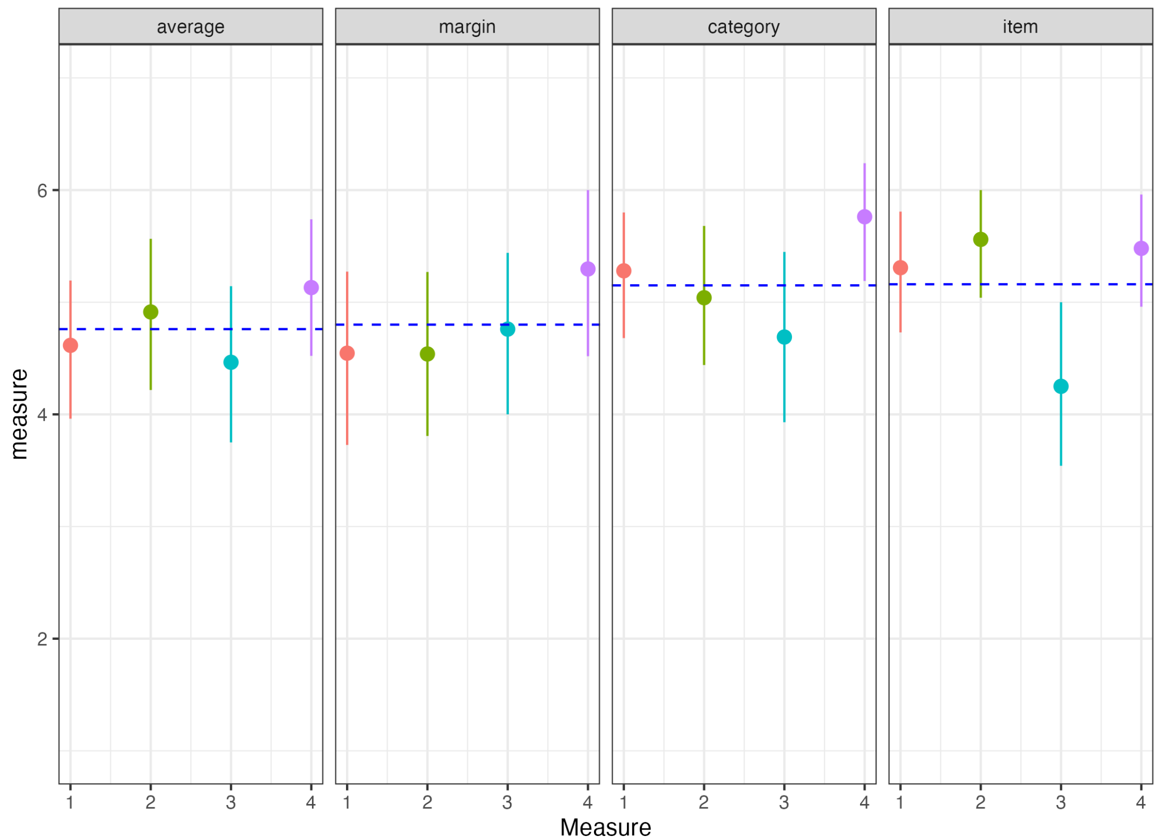
Monthly finances. Participants reported a median monthly take-home pay of $2500 (*M* = $4588, *SD* = $9225). Of this, the median amount dedicated to essential expenses was $1550 (*M* = $6120, *SD* = $35559). Both distributions were quite skewed, as suggested by the comparison of median and mean, alongside the reported standard deviations (data, code, plots, and additional distributional information provided in our materials at: https://researchbox.org/353&PEER_REVIEW_passcode=MIJYNO). For checks of robustness, we identify any observations farther than two standard deviations from the mean (+/- 2 SD). We subsequently reconsider our analyses with these observations removed.

Recall, participants allocated discretionary funds using the pie chart budgeting tool. To calculate discretionary funds, we subtracted expenses from take-home pay. At the individual level, the median discretionary amount was $650. A total of 15/100 participants had discretionary levels below $100. We asked participants with sub-$100 discretionary funds to continue the exercise as though they had $100 available. Therefore, the mean amount of discretionary funds ($1566; *SD* = $2850) is dependent on this specific approach. A total of 3 participants had actual discretionary funds equal to $100, yielding the reported 18/100 participants who allocated $100 of funds, as reported in the manuscript.

Allocations. Though not central to the question of how budgeters perceive value (in terms of either average value or marginal value), we observe the self-reported discretionary budgets. On average, participants allocated 39% of their discretionary funds to groceries, 14% to dining out, 21% to entertainment, and 16% to clothing. The remaining 10% was allocated across self-generated categories (used by 23/100 participants). Additional descriptions and analyses of budgeters' allocations are provided in our ResearchBox materials.

Self-reported focus measures. The complete descriptive results pertaining to the self-reported measures for average value, marginal value, category focus, and item focus are presented in figure B.1. As previously discussed, there was no effect of the question ordering (the two counterbalanced factors) on these measures. Furthermore, these measures (as well as the main reported findings in the manuscript) are robust to excluding participants with extreme incomes or expenses (+/- 2 SD) and the participants with imputed $100 discretionary spending. The associated plots and tests are available in our online materials.

FIGURE B.1
SELF-REPORTED FOCUS MEASURES ACROSS ALL QUESTION VARIANTS



Note—Self-reported focus on the dimensions of average value , marginal value, category focus, and item focus across all four question variants. Higher scores indicate an increased focus on the specific dimension of value. Error bars are 95% confidence intervals. The solid blue lines independently depict the mean average value and marginal value. The question variant refers to the "variant id" in table B.1.

WEB APPENDIX C: SUPPLEMENT TO STUDY 2

We preregistered a set of independent regression models with the intent of presenting our findings graphically.

$$ALLOCATE_{k,i} = b_{0_k} + b_{1_k}ABOVE_{k,i} + b_{2_k}MARGIN_{k,i} + b_{3_k}BELOW_{k,i} \qquad \text{(EQ 1)}$$

For the range of data (*k* = [3, 13]), this amounts to the following independent regression equations, which are estimated in table C.1B and plotted in figure C.1.

$$ALLOCATE_{3,i} = b_{0_3} + b_{1_3}ABOVE_{3,i} + b_{2_3}MARGIN_{3,i} + b_{3_3}BELOW_{3,i}$$

$$ALLOCATE_{4,i} = b_{0_4} + b_{1_4}ABOVE_{4,i} + b_{2_4}MARGIN_{4,i} + b_{3_4}BELOW_{4,i}$$

…

$$ALLOCATE_{13,i} = b_{0_{13}} + b_{1_{13}}ABOVE_{13,i} + b_{2_{13}}MARGIN_{13,i} + b_{3_{13}}BELOW_{13,i}$$

We note that the preregistered approach (EQ 1) can be equivalently presented as a set of separate equations or a single a nested model, in which there are no main effects, but rather a set of simple effects (for each level of *k*). The coefficient estimates are equivalent, though pooling the data and accounting for non-independence of observations results in similar but not identical standard errors.

For ease of explication, we deviate from our preregistered plan by unnesting the estimates of ABOVE, MARGIN, and BELOW in order to produce estimated main effects for each measure of value, where FE_k represents fixed effects for each rank. Specifically, we consider the model presented as EQ 2 (with cluster-robust standard errors).

$$ALLOCATE_{k,i} = b_0 + b_1 ABOVE_{k,i} + b_2 MARGIN_{k,i} + b_3 BELOW_{k,i} + FE_k \qquad \text{(EQ 2)}$$

These results are presented in table C.0 and in the main text. As a technical note, the main effects in the alternate model (EQ 2) may be roughly considered as the weighted average of the simple slopes across nested levels in the original model (EQ 1). Therefore, we present the alternate model (EQ 2) in the manuscript for the simplicity of reporting a single set of main effects (ABOVE, MARGIN, BELOW).

TABLE C.0
ESTIMATES FROM UNNESTED MODEL IN MANUSCRIPT

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
a	0.0008	0.0087	0.0946	0.9248	-0.0165	0.0181	147.8891
m	0.0197	0.0033	5.9286	0.0000	0.0132	0.0263	280.2366
b	0.0351	0.0063	5.5580	0.0000	0.0227	0.0476	201.5317

Note—Main effects from a single unnested model for ABOVE (“a”), MARGIN (“m”), and BELOW (“b”), estimated using *lm_robust()* with cluster-robust standard errors (at the participant level) and rank fixed effects (estimatr package in R).

C.1: Preregistered Analysis

Nested Model

$$ALLOCATE_{k,i} = b_{0_k} + b_{1_k}ABOVE_{k,i} + b_{2_k}MARGIN_{k,i} + b_{3_k}BELOW_{k,i}$$

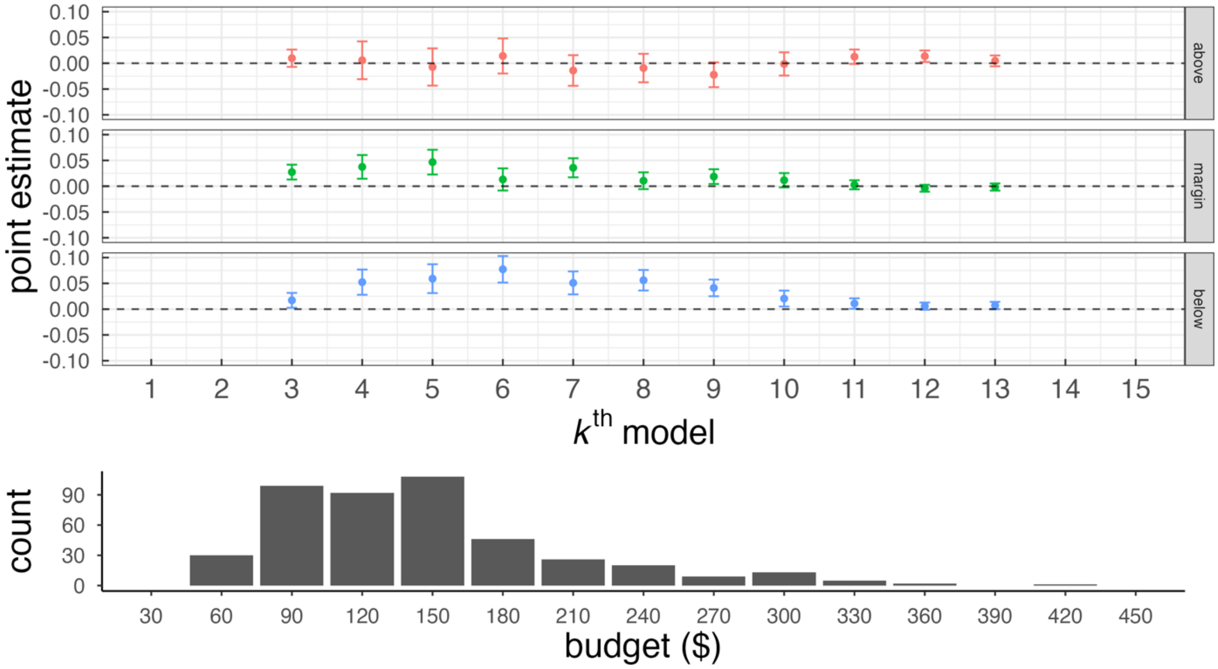
TABLE C.1A
VARIABLE NAME AND CONSTRUCTION

Variable Name	**Variable Construction**
ALLOCATE	Indicator that participant i allocates enough funds for at least k activities [0/1]
ABOVE	Mean rated value of all considered options ranked better than k for participant i
MARGIN	Rated value of the k^{th} option for participant i
BELOW	Mean rated value of all considered options ranked worse than k for participant i

The preregistered plan was to have ABOVE, MARGIN, and BELOW are nested under each level of *k*, in which the coefficients of interest can be presented as either (i) the set of simple slopes (in a single nested model with ABOVE, MARGIN, and BELOW interacting with each level of *k*) or (ii) independent *k*-level regressions. For clarity, we present the results following this latter approach. The point estimates for ABOVE, MARGIN, and BELOW are plotted in figure C.1 and the corresponding regression results are in table C.1B.

FIGURE C.1

PREREGISTERED RESULTS AND BUDGET SIZE



Note—The top panel plots the point estimates for ABOVE, MARGIN, and BELOW across independent regressions (at various levels of *k*). Error bars are 95% confidence intervals. The bottom panel plots the frequencies of participants budgeting a given amount, in dollars.

TABLE C.1B

PREREGISTERED RESULTS

	Dependent variable:										
	dv3	dv4	dv5	dv6	dv7	dv8	dv9	dv10	dv11	dv12	dv13
ABOVE	0.01	0.01	−0.01	0.01	−0.01	−0.01	-0.02^{+}	−0.001	0.01^{+}	0.01^{*}	0.005
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
MARGIN	0.03^{***}	0.04^{**}	0.05^{***}	0.01	0.04^{***}	0.01	0.02^{*}	0.01	0.003	−0.004	−0.002
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.003)	(0.004)
BELOW	0.02^{*}	0.05^{***}	0.06^{***}	0.08^{***}	0.05^{***}	0.06^{***}	0.04^{***}	0.02^{**}	0.01^{*}	0.01	0.01^{+}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.004)
Constant	0.54^{***}	0.14	−0.01	-0.25^{+}	−0.13	−0.08	−0.002	−0.06	-0.12^{*}	-0.08^{*}	−0.04
	(0.08)	(0.14)	(0.15)	(0.13)	(0.11)	(0.10)	(0.08)	(0.07)	(0.05)	(0.03)	(0.03)
Observations	446	438	434	424	409	385	349	305	272	227	175

Note: +p<0.10;*p<0.05;**p<0.01;***p<0.001

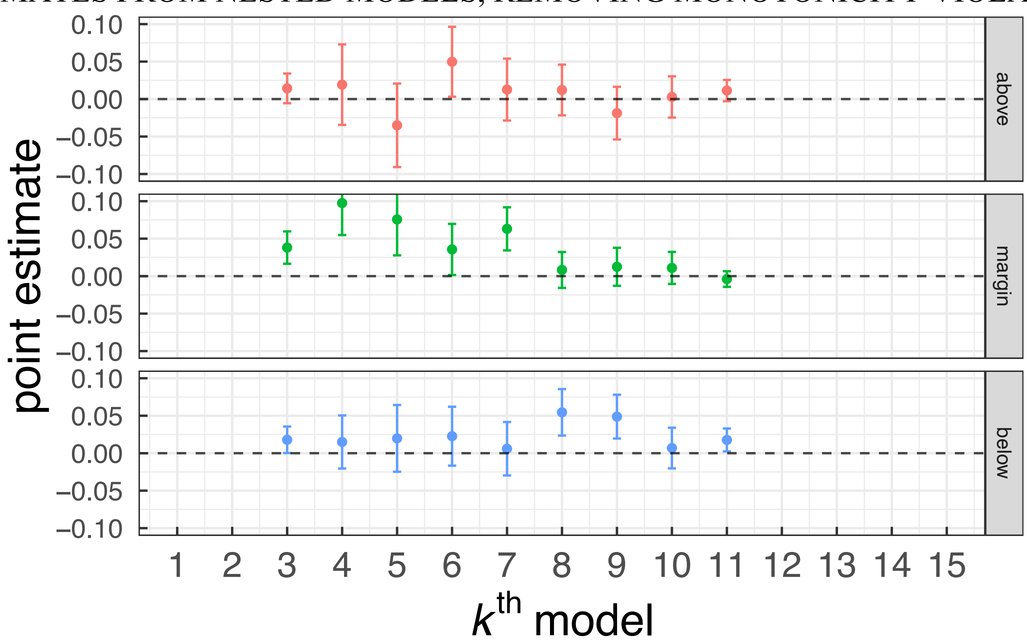
C.2: Robustness: Removing monotonicity violations

Nested Model

The model in analysis C.2 is identical to the primary model, except we exclude any observations in which the marginal value (the k^{th} ranked option) is rated lower than any worse-ranked alternative. We provide both the nested (figure C.2, table C.2A) and unnested results (table C.2B). As a technical note, as a result of excluding observations that are not at least weakly monotonically decreasing (when considering the rated value-to-rank relationship), there is no remaining variation in the dependent variable (allocating sufficient funds for a given level of k) when $k = 12, 13$. For this reason, we cannot create estimates for ABOVE, MARGIN, and BELOW when $k > 11$.

FIGURE C.2

ESTIMATES FROM NESTED MODELS, REMOVING MONOTONICITY VIOLATIONS



Note—This plot portrays point estimates of the independent regressions in a manner identical to figure C.2, except: (1) regressions are estimated after removing any monotonicity violations and (2) models for $k = 12, 13$ are no longer estimable due to data loss.

TABLE C.2A
ESTIMATES FROM NESTED MODELS AFTER
REMOVING MONOTONICITY VIOLATIONS

	Dependent variable:								
	dv3	dv4	dv5	dv6	dv7	dv8	dv9	dv10	dv11
ABOVE	0.01	0.02	−0.04	0.05**	0.01	0.01	−0.02	0.003	0.01
	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
MARGIN	0.04***	0.10***	0.08***	0.04**	0.06***	0.01	0.01	0.01	−0.004
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BELOW	0.02**	0.01	0.02	0.02	0.01	0.05***	0.05***	0.01	0.02**
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Constant	0.39***	−0.34*	0.15	−0.52***	−0.37**	−0.21*	−0.002	−0.05	−0.09**
	(0.09)	(0.21)	(0.21)	(0.19)	(0.15)	(0.11)	(0.11)	(0.09)	(0.05)
Observations	300	250	225	237	229	225	198	166	163
Note:						+p<0.10;*p<0.05;**p<0.01;***p<0.001			

Unnested Model

Unnesting the model to consider main effects:

TABLE C.2B
ESTIMATES FROM UNNESTED MODEL AFTER
REMOVING MONOTONICITY VIOLATIONS

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
a	0.0117	0.0092	1.2746	0.2048	-0.0065	0.0299	128.9565
m	0.0313	0.0061	5.1270	0.0000	0.0193	0.0434	176.8822
b	0.0244	0.0085	2.8871	0.0043	0.0077	0.0411	193.0599

Note—Main effects from a single unnested model for ABOVE (“a”), MARGIN (“m”), and BELOW (“b”), estimated using *lm_robust()* with cluster-robust standard errors (at the participant level) and rank fixed effects (estimatr package in R).

C.3: Including non-considered activities (with zero values)

Nested Model

$$ALLOCATE_{k,i} = b_{0_k} + b_{1_k}ABOVE_{k,i} + b_{2_k}MARGIN_{k,i} + b_{3_k}BELOW_{k,i}$$

A key feature of the design of study 2 was to allow participants to identify the considered set (vs. non-considered set). Our preregistered approach, manuscript approach, and the previously discussed robustness checks are all constrained to the values of considered activities. As an additional check, we can also incorporate the non-considered activities into the model with an imputed value of 0.[1] As can be seen in table C.3A, this imputation affects only the construction of BELOW. Specifically, the average value of activities ranked worse than *k* (what is captured by BELOW) will decrease for participants who considered < 15 activities, as this measure will be dragged down by the inclusion of 0 values. We present results using the nested model (figure C.3, table C.3B) and unnested model (table C.3C).

TABLE C.3A
VARIABLE NAME AND CONSTRUCTION

Variable Name	Variable Construction
ALLOCATE	Indicator that participant *i* allocates enough funds for at least *k* activities [0/1]
ABOVE	Mean rated value of all considered options ranked better than *k* for participant *i*
MARGIN	Rated value of the k^{th} option for participant *i*
BELOW	Mean rated value of all options ranked worse than *k* for participant *i*, including imputed 0 values for non-considered options

[1] Of course, there may be different, non-linear functional forms to assign imputed values to non-considered items; however, this is beyond the scope of our analysis.

FIGURE C.3
ESTIMATES FROM NESTED MODELS AFTER IMPUTING 0 VALUES
FOR NON-CONSIDERED ACTIVITIES

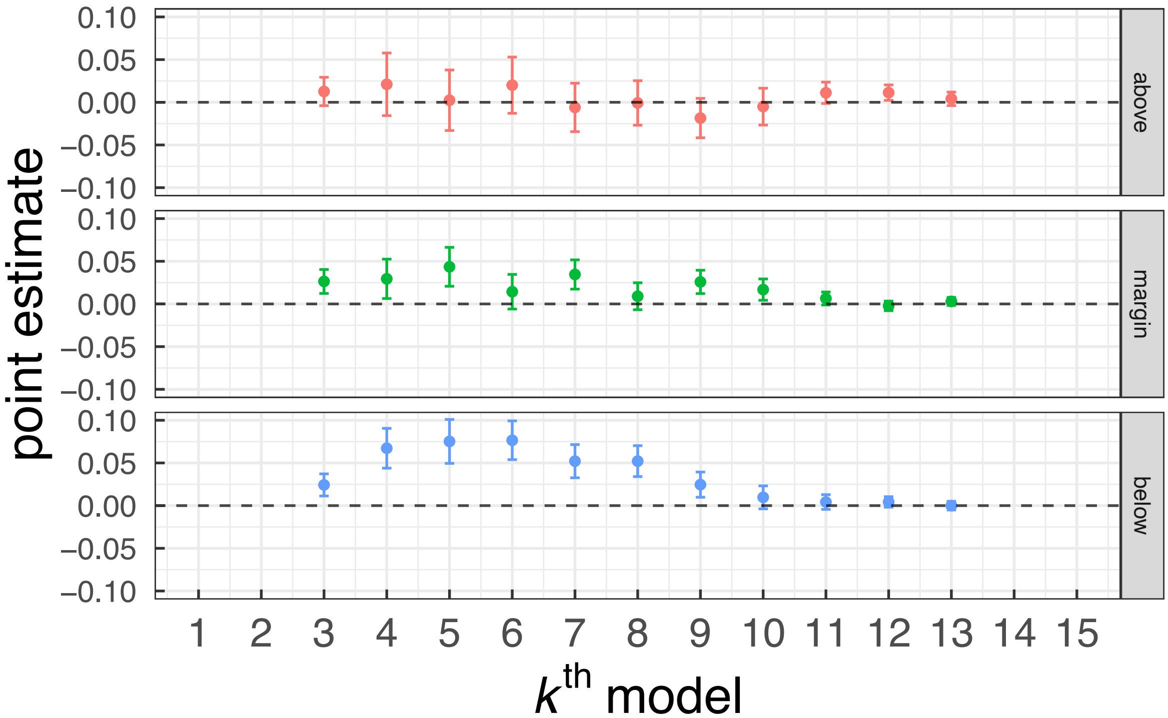


TABLE C.3B
ESTIMATES FROM NESTED MODELS AFTER IMPUTING 0 VALUES
FOR NON-CONSIDERED ACTIVITIES

	Dependent variable:										
	dv3	dv4	dv5	dv6	dv7	dv8	dv9	dv10	dv11	dv12	dv13
ABOVE	0.01	0.02	0.002	0.02	−0.01	−0.001	−0.02	−0.01	0.01^{+}	0.01^{*}	0.004
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)	(0.004)
MARGIN	0.03^{***}	0.03^{*}	0.04^{***}	0.01	0.03^{***}	0.01	0.03^{***}	0.02^{**}	0.01^{+}	−0.002	0.003
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.003)	(0.002)
BELOW	0.02^{***}	0.07^{***}	0.08^{***}	0.08^{***}	0.05^{***}	0.05^{***}	0.02^{**}	0.01	0.004	0.004	−0.0000
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.003)	(0.003)
Constant	0.52^{***}	0.08	−0.04	−0.20	−0.12	−0.05	0.03	−0.004	-0.09^{*}	-0.07^{*}	−0.03
	(0.08)	(0.14)	(0.14)	(0.13)	(0.11)	(0.09)	(0.08)	(0.07)	(0.04)	(0.03)	(0.02)
Observations	449	446	438	434	424	409	385	349	305	272	227
Note:											+p<0.10;*p<0.05;**p<0.01;***p<0.001

Unnested Model

Unnesting the model to consider main effects:

TABLE C.3C
ESTIMATES FROM UNNESTED MODEL AFTER IMPUTING 0 VALUES FOR NON-CONSIDERED ACTIVITIES

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
a	0.0043	0.0082	0.5222	0.6023	-0.0119	0.0205	158.7981
m	0.0188	0.0032	5.7725	0.0000	0.0124	0.0251	290.9755
b	0.0364	0.0056	6.4431	0.0000	0.0252	0.0475	160.7004

Note—Main effects from a single unnested model for ABOVE (“a”), MARGIN (“m”), and BELOW (“b”), estimated using *lm_robust()* with cluster-robust standard errors (at the participant level) and rank fixed effects (estimatr package in R).

C.4: Average of all considered options

Nested Model

$$ALLOCATE_{k,i} = b_{0_k} + b_{1_k}AVERAGE_{k,i} + b_{2_k}MARGIN_{k,i}$$

The primary hypothesis we test is whether budget allocations are sensitive to a category's average value, above and beyond marginal value (H1). Study 2 benefits from our ability to separate the average value of activities ranked better and worse than the marginal good (corresponding to ABOVE and BELOW) to understand what region of the value distribution may be driving the sensitivity to average value. However, it is still important to understand whether participants are sensitive to the simple average value, calculated from all considered options, aside from the marginal activity. Results from the nested model are presented in figure C.4 and table C.4B, and the main effects from the unnested model are given by table C.4C.

TABLE C.4A
VARIABLE NAME AND CONSTRUCTION

Variable Name	Variable Construction
ALLOCATE	Indicator that participant *i* allocates enough funds for at least *k* activities [0/1]
AVERAGE	Mean rated value of all considered options except the k^{th} option for participant *i*
MARGIN	Rated value of the k^{th} option for participant *i*

FIGURE C.4

ESTIMATES FROM NESTED MODELS USING A SINGLE MEASURE OF AVERAGE

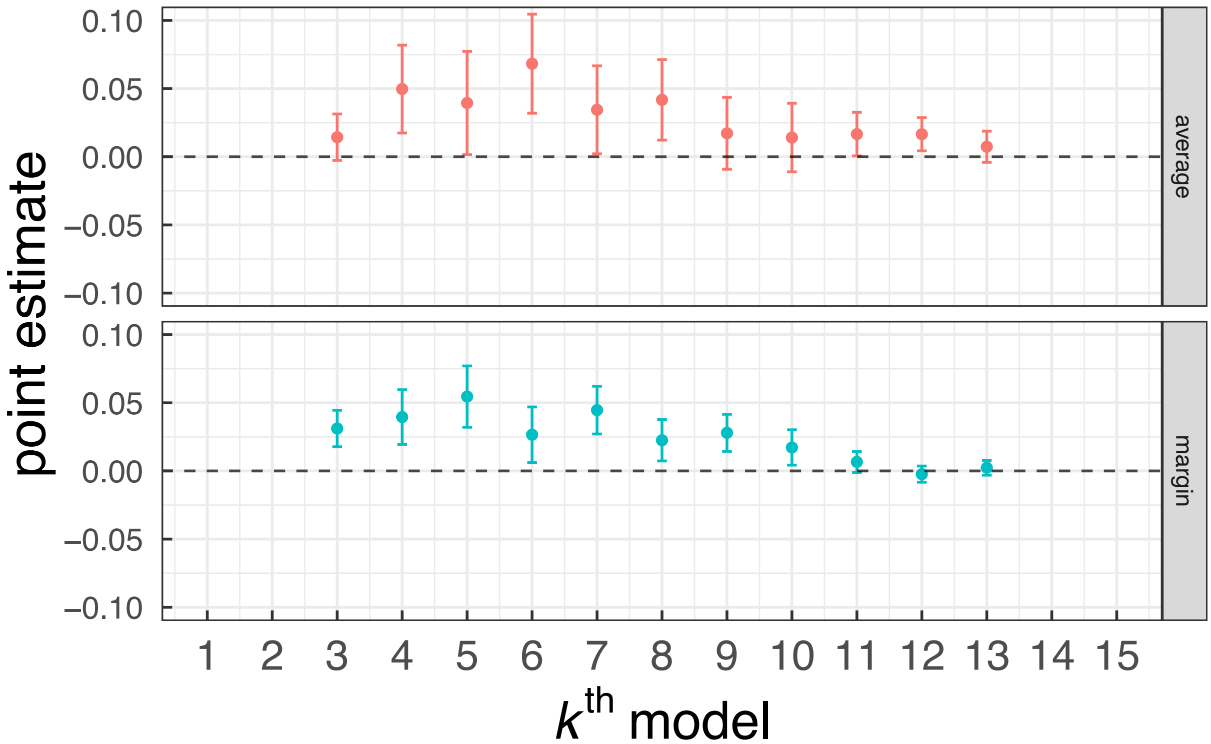


TABLE C.4B

ESTIMATES FROM NESTED MODELS USING A SINGLE MEASURE OF AVERAGE

	Dependent variable:										
	dv3	dv4	dv5	dv6	dv7	dv8	dv9	dv10	dv11	dv12	dv13
AVERAGE	0.01	0.05***	0.04**	0.07***	0.03**	0.04***	0.02	0.01	0.02**	0.02***	0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
MARGIN	0.03***	0.04***	0.05***	0.03**	0.04***	0.02***	0.03***	0.02***	0.01*	−0.002	0.002
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.003)	(0.003)
Constant	0.60***	0.13	−0.08	−0.30***	−0.28***	−0.25***	−0.16**	−0.11	−0.11***	−0.08**	−0.05
	(0.07)	(0.10)	(0.11)	(0.10)	(0.09)	(0.08)	(0.07)	(0.07)	(0.04)	(0.03)	(0.03)
Observations	446	438	434	424	409	385	349	305	272	227	175

Note: +p<0.10;*p<0.05;**p<0.01;***p<0.001

Unnested Model

Unnesting the model to consider main effects:

TABLE C.4C
ESTIMATES FROM UNNESTED MODEL USING A SINGLE MEASURE OF AVERAGE

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
m	0.0264	0.0033	7.8954	0.0000	0.0199	0.0330	283.6958
avg	0.0319	0.0082	3.9059	0.0001	0.0158	0.0480	236.3242

Note—Main effects from a single unnested model for MARGIN (“m”) and the overall average of considered activities (“avg”), estimated using *lm_robust()* with cluster-robust standard errors (at the participant level) and rank fixed effects (estimatr package in R).

C.5: Substituting Adjacent Values for Averages

Nested Model

$$ALLOCATE_{k,i} = b_{0_k} + b_{1_k}ABOVE_{k,i} + b_{2_k}MARGIN_{k,i} + b_{3_k}BELOW_{k,i}$$

While our discussion of the literature of ensemble perception suggests people extract average representations from the considered set, we additionally test whether this sensitivity is captured by the valuations of activities with the most similar values to the marginal item. Specifically, we construct ABOVE as the single value of the activity ranked directly better and BELOW as the single value of the activity ranked directly worse. Results for the nested model are presented in figure C.5 and table C.5B, and the main effects from the unnested model are given in table C.5C.

TABLE C.5A
VARIABLE NAME AND CONSTRUCTION

Variable Name	Variable Construction
ALLOCATE	Indicator that participant *i* allocates enough funds for at least *k* activities [0/1]
ABOVE	Rated value of option ranked just better than *k* for participant *i*
MARGIN	Rated value of the k^{th} option for participant *i*
BELOW	Rated value of option ranked just worse than *k* for participant *i*

FIGURE C.5

ESTIMATES FROM NESTED MODELS USING ADJACENT VALUES FOR AVERAGES

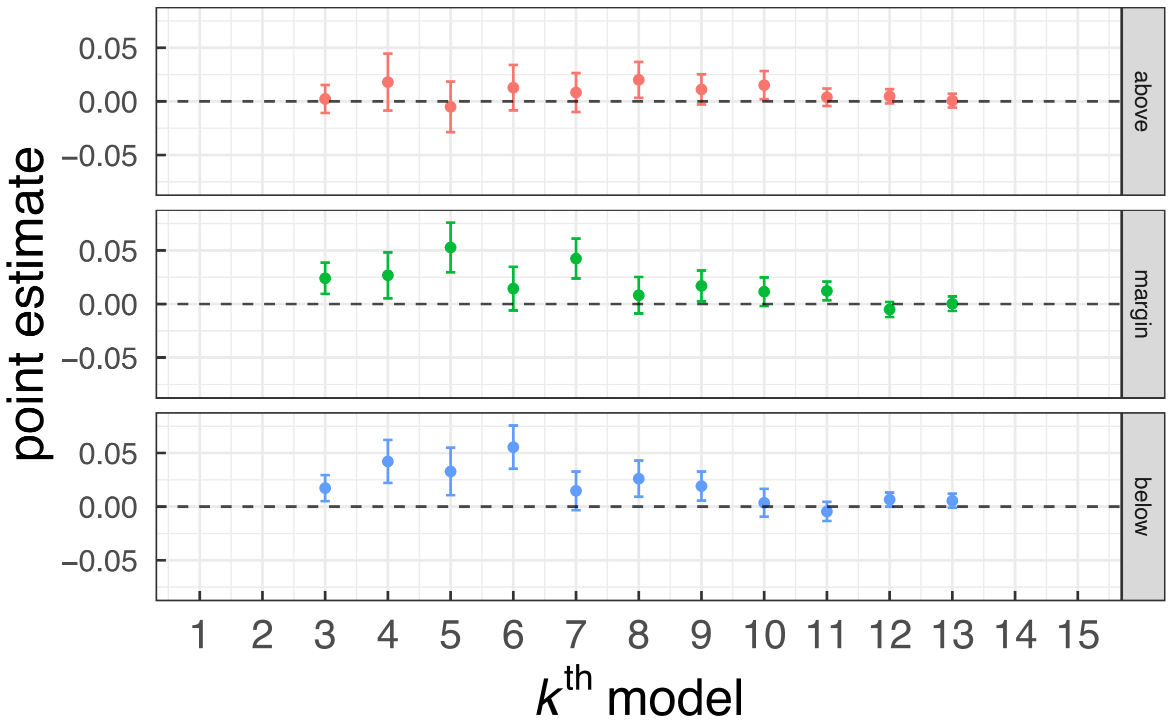


TABLE C.5B

ESTIMATES FROM NESTED MODELS USING ADJACENT VALUES FOR AVERAGES

	Dependent variable:										
	dv3	dv4	dv5	dv6	dv7	dv8	dv9	dv10	dv11	dv12	dv13
ABOVE	0.002	0.02	−0.01	0.01	0.01	0.02**	0.01	0.02**	0.004	0.005	0.001
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.003)	(0.003)
MARGIN	0.02***	0.03**	0.05***	0.01	0.04***	0.01	0.02**	0.01*	0.01***	−0.01	0.0003
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.004)	(0.003)
BELOW	0.02***	0.04***	0.03***	0.06***	0.01	0.03***	0.02***	0.004	−0.004	0.01*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)	(0.003)	(0.003)
Constant	0.59***	0.09	−0.003	−0.21***	−0.19***	−0.16***	−0.15***	−0.09***	−0.04*	−0.02	−0.02
	(0.06)	(0.11)	(0.10)	(0.08)	(0.06)	(0.05)	(0.04)	(0.03)	(0.02)	(0.01)	(0.01)
Observations	446	438	434	424	409	385	349	305	272	227	175

Note: +p<0.10;*p<0.05;**p<0.01;***p<0.001

Unnested Model

Unnesting the model to consider main effects:

TABLE C.5C
ESTIMATES FROM UNNESTED MODEL USING ADJACENT VALUES FOR AVERAGES

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df
a	0.0056	0.0029	1.9293	0.0547	-0.0001	0.0112	278.3154
m	0.0200	0.0027	7.3482	0.0000	0.0146	0.0253	258.0804
b	0.0229	0.0031	7.3432	0.0000	0.0167	0.0290	288.1172

Note—Main effects from a single unnested model for the single value of the activity just ABOVE the margin (“a”), MARGIN (“m”), and the single value of the activity just BELOW the margin (“b”), estimated using *lm_robust()* with cluster-robust standard errors (at the participant level) and rank fixed effects (estimatr package in R).

C.6: Allocation Predicts Spending

Allocation predicts spending. We examine the hypothetical spending decisions by considering the 435/451 participants who considered more activities than their budget would allow (e.g., considered and provided value ratings to 12 activities, and set a $150 budget to accommodate up to 5 activities).[2] For these participants, activity value is a strong predictor of whether a given activity is selected for purchase, controlling for budget level and clustering standard errors at the participant level ($t(6764) = 56.31$, $p < .001$). Considering individual participants, 48% (208/435) purchased the optimal set of items (the highest ranked items, given their budget constraint) and the average proportion of optimal purchases across subjects was 84%. We take these findings as evidence that participants had stable and meaningful preferences (in terms of both values and ranks) that informed hypothetical purchase decisions in our paradigm.

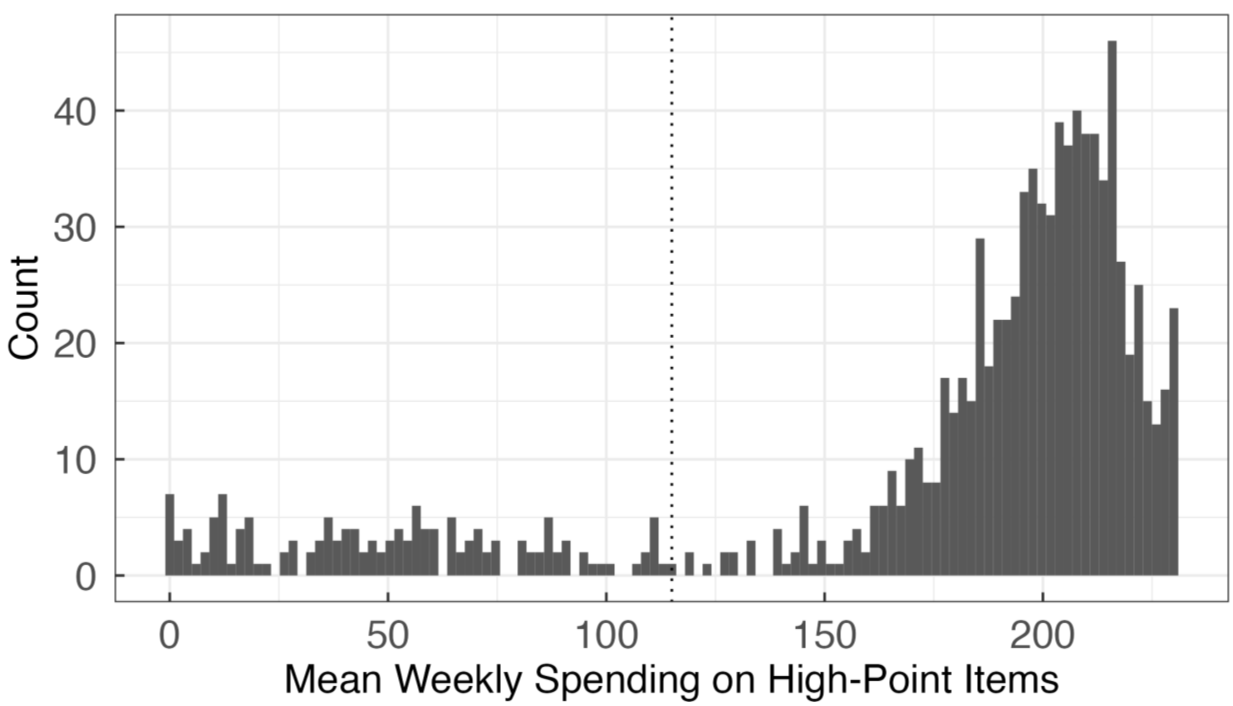
[2] For participants who considered *fewer* activities than allocated for, the purchased set is mechanically guaranteed to be the considered set.

WEB APPENDIX D: SUPPLEMENT TO STUDY 3

Exclusions and Comprehension Checks

Exclusions. Figure D.1 depicts the distribution of spending on high-point purchases used for exclusion in study 3. Participants who bought less than half of the available high-point purchases (those to the left of the dashed line) were preregistered to be excluded, as these participants were likely inattentive or misunderstood the game.

FIGURE D.1
DISTRIBUTION OF MEAN WEEKLY HIGH-POINT SPENDING, USED FOR EXCLUSION



Note— Distribution of spending on high-point items in study 3. The mass of the data lay well above 50% (dotted line). Purchasing less than 50% of these high-point items (participants to the left of the dotted line) is outside the range of typical behavior and is taken to indicate inattention or misunderstanding of the task.

Comprehension. Among the included participants, performance on comprehension questions was quite good. Correct response rates were 92%, 97%, 81%, 98%, and 87%, corresponding to the five sequential questions (below). After answering each question, participants were provided feedback about whether their response was correct or incorrect and given an explanation.

- *Q1*: "The goal of this game is to collect as many____as possible": 92% correctly identified "points" from a list of three options ("points," "entertainment purchases," or "dining purchases").

- *Q2a*: "Which of the following is true about the budgets for dining and entertainment purchases?": 97% correctly identified "Budgets may help plan my purchases, though I am not required to follow them" rather than "budgets must be followed exactly." (This question was only asked of participants in the budget condition.)

Questions Q2b-Q4 were True or False.

- *Q2b:* "During a given week, you may make 23 or fewer purchases." 81% correctly answered "True."

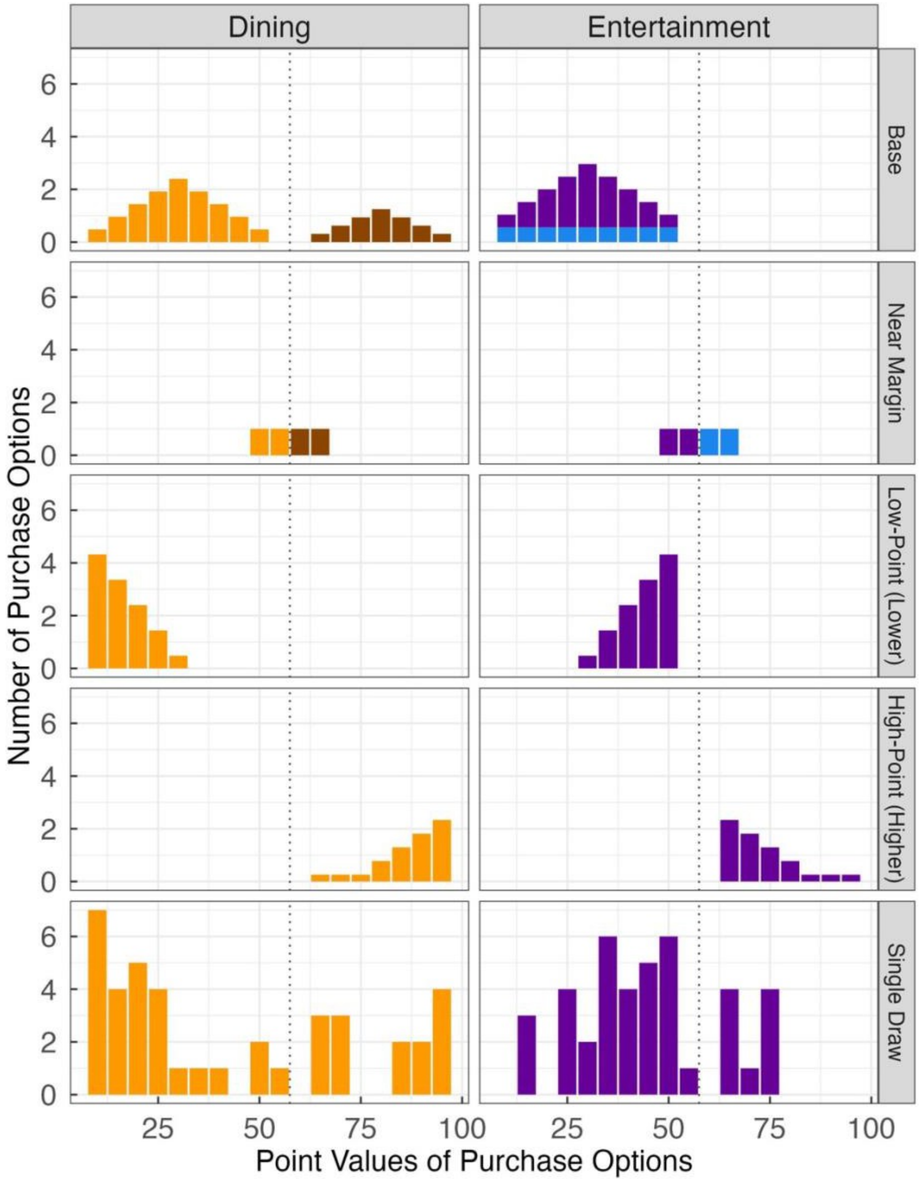
- *Q3:* "Any unspent money will carry over to the following week": 98% correctly answered "False."

- *Q4:* “You will have the opportunity to earn a bonus during both the five-week practice round and the five-week game.”: 87% correctly answered “False.”

Construction of Distributions

Figure D.2 depicts the theoretical distributions from which items were drawn in study 3. The first row indicates the common portions of the dining and entertainment distributions used in all conditions. The second row indicates the distribution for the two best options from the low-point region of the distribution and the two worst options from the high-point region of the distribution. The third row depicts the low-point region of the distribution when dining is low and entertainment is high; these distributions would be swapped in the condition where dining is high and entertainment is low. The fourth row depicts the high-point region of the distribution when dining is high and entertainment is low; these distributions would be swapped in the other high-value condition. By drawing items from these distributions, there were always exactly 14 dining options worth at least 60 points and there were always exactly 9 entertainment options worth at least 60 points, but the category average values systematically varied by condition. Finally, the bottom row depicts a sample draw from the theoretical distributions in the prior four rows. This single draw is typical of what a participant in that condition may have been presented with.

FIGURE D.2
THEORETICAL POINT DISTRIBUTIONS IN STUDY 3



Note—Distributions from which items were drawn in study 3. The dotted vertical line represents the split between the low-point region of the distribution (point values less than 60) and the high-point region of the distribution (point values of 60 or more). Row 1 ensured that possible points did not systematically differ across conditions. 5 dining options were drawn from the brown distribution; 12 dining items were drawn from the orange distribution; 12 entertainment items were drawn from the purple distribution; 5 entertainment options were drawn from the blue distribution. Row 2 ensured deviations of up to 2 items from the value- maximizing bundle would lead to symmetric outcomes. 2 items were drawn from each of the orange, brown, purple, and blue distributions. These options were the best low-value options (either 50 or 55) and the worst high-value options (either 60 or 65) available. Rows 3 and 4 depict the manipulation of the low-value and high-value parts of the distributions. Row 5 depicts a single sample draw a participant may have seen.

Additional Results

For completeness, we provide the full regression output for the analyses accompanying each hypothesis. Furthermore, we include both the dining share measure reported in the manuscript (for ease of explication), as well as the difference measure (as preregistered)

First analysis (H1). We present the regression output for the full preregistered model in table D.1A. As discussed in the manuscript, this model was constrained to those in the budgeting condition and regressed the dining share of allocation on both the dining average, the low-point dining average, and the interaction. We include an analysis over the full data (without exclusions) for completeness, as well as using the preregistered difference measure in columns 3 and 4.

TABLE D.1A

REGRESSION RESULTS TO ACCOMPANY TEST OF H1

	Dependent variable:			
	Dining Share		Difference Measure	
	With Exclusions	Without Exclusions	With Exclusions	Without Exclusions
	(1)	(2)	(3)	(4)
Dining average	2.606***	2.055***	11.990***	9.454***
	p = 0.00000	p = 0.0001	p = 0.00000	p = 0.0001
Low-point dining average	0.832+	1.286*	3.829+	5.916*
	p = 0.089	p = 0.012	p = 0.089	p = 0.012
Din. avg. x low din. avg.	0.065	0.012	0.299	0.057
	p = 0.895	p = 0.981	p = 0.895	p = 0.981
Constant	54.983***	55.294***	22.922***	24.354***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	394	478	394	478
R^2	0.075	0.046	0.075	0.046
Adjusted R^2	0.068	0.040	0.068	0.040
Residual Std. Error	9.675 (df = 390)	11.114 (df = 474)	44.503 (df = 390)	51.124 (df = 474)

Note: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Note—Dependent variable is the dining share of allocation (cols 1-2) and the difference in dining (3-4). Dining average refers to the manipulation of average category values in the high-point region of the distribution (+1 = dining high, -1 = dining low). Low-point dining average refers to the manipulation of average category values in the low-point region of the distribution (+1 = dining high, -1 = dining low). Columns 1 and 3 apply the preregistered exclusions and columns 2 and 4 considers the full data, without exclusions.

Additional exploratory analysis: sensitivity to overall average value. While our experimental design allows for the careful estimation of sensitive to average value in the high-point and low-point regions of the distribution, we can also approximate the sensitivity to overall average value. A simple way to conduct this approximation is to reconceptualize the original 2 (dining average: high, low) x 2 (low-point dining average: high, low) as a 3-condition, between-subjects design (high averages in both regions, mixed high and low, low averages in both regions). We reconsider the test of H1 after constructing a new variable for overall average, reflecting this condition assignment (+1 = both high averages, 0 = mixed high and low averages, -1 both low averages). Results are presented in Table D.1B.

TABLE D.1B
REGRESSION RESULTS TO ACCOMPANY TEST OF H1

	Dependent variable:
	Dining share
Overall average	3.439***
	p = 0.00001
Constant	55.015***
	p = 0.000
Observations	394
R^2	0.059
Adjusted R^2	0.057
Residual Std. Error	9.731 (df = 392)
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Note—Exploratory analysis suggests budgeters are sensitive to the overall average value, constructed from original condition assignment across the 2x2 design.

Second analysis (H2). As discussed in the manuscript, the test of H2 uses the same model as that to test H1 (including being constrained to only those in the budget condition), with one difference. Whereas H1 considers the dining share of *allocation* as the dependent variable, H2 considers the dining share of *spending*. For completeness, we present results with and without preregistered exclusions, as well as using both the proportional dining share measure (from the manuscript) as well as the preregistered difference measure (table D.2).

TABLE D.2
REGRESSION RESULTS TO ACCOMPANY TEST OF H2

	Dependent variable:			
	Dining Share of Spending		Difference Measure	
	With Exclusions	Without Exclusions	With Exclusions	Without Exclusions
	(1)	(2)	(3)	(4)
Dining average	2.431***	2.232***	10.863***	9.710***
	p = 0.000	p = 0.00000	p = 0.000	p = 0.000
Low-point dining average	1.241**	0.948*	5.617***	4.912**
	p = 0.002	p = 0.022	p = 0.001	p = 0.004
Din. avg. x low din. avg.	−0.182	−0.381	−0.712	−1.830
	p = 0.627	p = 0.354	p = 0.668	p = 0.271
Constant	57.225***	57.272***	32.512***	30.502***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	394	478	394	478
R^2	0.120	0.071	0.123	0.086
Adjusted R^2	0.114	0.065	0.116	0.080
Residual Std. Error	7.429 (df = 390)	8.975 (df = 474)	32.858 (df = 390)	36.256 (df = 474)
Note:			+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Note—Dependent variable is the dining share of allocation (cols 1-2) and the difference in dining (3-4). Dining average refers to the manipulation of average category values in the high-point region of the distribution (+1 = dining high, -1 = dining low). Low-point dining average refers to the manipulation of average category values in the low-point region of the distribution (+1 = dining high, -1 = dining low). Columns 1 and 3 apply the preregistered exclusions and columns 2 and 4 considers the full data, without exclusions.

Third analysis (H3). Though not preregistered, we can consider the dining share of *allocation* among budgeters and the dining share of *spending* among non-budgeters as a single dependent measure. Specifically, the dining share reflects [dining dollars / total allocated dollars] x 100% for budgeters and [dining dollars / total spent dollars] x 100% for non-budgeters. This allows us to include all participants (those assigned to both conditions) in the analysis. We regress this dining share on the dining average (+1 = high, -1 = low), the low-point dining average (+1 = high, -1 = low), the budget condition assignment (+1 = budgeting, -1 = non-budgeting), and all two- and three-way interactions. Full results are presented in table D.3. The variable of interest is the dining average-by-budget interaction. Columns 3 and 4 include the difference measure.

TABLE D.3

REGRESSION RESULTS TO ACCOMPANY TEST OF H3

	Dependent variable:			
	Dining Share		Difference Measure	
	With Exclusions	Without Exclusions	With Exclusions	Without Exclusions
	(1)	(2)	(3)	(4)
Dining average	1.977***	1.695***	8.979***	7.536***
	p = 0.000	p = 0.00000	p = 0.000	p = 0.000
Low-point dining average	1.336***	1.167***	6.049***	5.626***
	p = 0.00000	p = 0.0002	p = 0.00000	p = 0.00002
Budget	−2.489***	−1.867***	−11.122***	−7.615***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Din. avg. x low din. avg.	0.055	0.078	0.264	0.129
	p = 0.823	p = 0.799	p = 0.815	p = 0.921
Din. avg. x budget	0.629*	0.361	3.010**	1.919
	p = 0.011	p = 0.242	p = 0.008	p = 0.139
Low din. avg. x budget	−0.503*	0.119	−2.220*	0.290
	p = 0.042	p = 0.700	p = 0.050	p = 0.823
Din. x low x budget	0.010	−0.066	0.034	−0.072
	p = 0.969	p = 0.831	p = 0.976	p = 0.956
Constant	57.472***	57.161***	34.045***	31.969***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	821	970	821	970
R^2	0.200	0.080	0.194	0.085
Adjusted R^2	0.193	0.073	0.188	0.079
Residual Std. Error	7.039 (df = 813)	9.573 (df = 962)	32.312 (df = 813)	40.253 (df = 962)
Note:			+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Note—Exploratory model to consider H3. Dependent variable is the dining share (cols 1-2) or dining difference (3-4) of allocation in the budget condition and the dining share of spending in the non-budget condition.

Fourth analysis (H4). The full regression output for our preregistered test of H4 is presented in table D.4. The dependent variable is the dining share of spending, and the focal preregistered variable is the dining average-by-budget interaction. Aside from the differences in the dependent variable, this model is identical to the model to test H3 (table D.3).

TABLE D.4
REGRESSION RESULTS TO ACCOMPANY TEST OF H4

	Dependent variable:			
	Dining Share		Difference Measure	
	With Exclusions	Without Exclusions	With Exclusions	Without Exclusions
	(1)	(2)	(3)	(4)
Dining average	1.890***	1.783***	8.416***	7.664***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Low-point dining average	1.540***	0.999***	6.943***	5.124***
	p = 0.000	p = 0.0003	p = 0.000	p = 0.00000
Budget	−1.368***	−0.878**	−6.327***	−4.541***
	p = 0.000	p = 0.002	p = 0.000	p = 0.00001
Din. avg. x low din. avg.	−0.068	−0.118	−0.241	−0.815
	p = 0.726	p = 0.661	p = 0.781	p = 0.418
Din. avg. x budget	0.542**	0.449+	2.447**	2.047*
	p = 0.006	p = 0.097	p = 0.005	p = 0.043
Low din. avg. x budget	−0.299	−0.050	−1.326	−0.212
	p = 0.126	p = 0.853	p = 0.126	p = 0.833
Din. x low x budget	−0.114	−0.263	−0.471	−1.015
	p = 0.560	p = 0.331	p = 0.586	p = 0.313
Constant	58.593***	58.150***	38.840***	35.043***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	821	970	821	970
R^2	0.206	0.069	0.212	0.103
Adjusted R^2	0.200	0.063	0.205	0.096
Residual Std. Error	5.579 (df = 813)	8.395 (df = 962)	24.737 (df = 813)	31.308 (df = 962)
Note:		+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001		

Note—Columns 1 and 3 apply the preregistered exclusions and columns 2 and 4 considers the full data, without exclusions.

Budgets predict spending in study 3: exploratory analysis. As an exploratory analysis, we regress each individual option's purchase decision (80 per week, for each of 5 weeks) on category, money remaining in the entertainment budget and money remaining in the dining budget, their interactions with category, and a rich set of controls. We include participant fixed effects to account for the fact that some participants routinely underspend on dining. We include week and day fixed effects to account for time trends. We include item-value fixed effects to reduce error. And we include history controls (the number of 95-point dining items seen, the number of 95-point entertainment items seen, the number of 90-point dining items seen,etc.) to account for expectations regarding category-specific remaining items. $10 remaining in the dining budget is associated with a 3.2 percentage point increase in the likelihood of purchasing a dining item but a 2.0 percentage point increase in the likelihood of purchasing an entertainment item; $10 remaining in the entertainment budget is associated with a 2.2 percentage point increase in the likelihood of purchasing a dining item but a 3.2 percentage point increase in the likelihood of purchasing an entertainment item. The difference between the differential effect of dining budget remaining on dining vs. entertainment spending and the differential effect of entertainment budget remaining on dining vs. entertainment spending is statistically significant ($t(65) = 7.78$, $p < .001$), indicating that funds are not treated as perfectly fungible.

Points and bonuses: exploratory analysis. As additional exploratory analyses, we consider both total points (accumulated throughout the incentivized five-week game) and total bonus payments as predicted by condition assignment. Bonuses were earned each week for scores of at least 1,560 points. Specifically, a participant earned an additional $0.01 bonus for every 10 points they scored above 1,550 points in a given week. Therefore, participants had five opportunities (five weeks) to earn non-zero bonuses, which accumulated to the final bonus payout (mean = \$0.75, median = \$0.80). The dining average condition and the low-point dining average condition did not affect points or bonuses (table D.5). Note that the distributions were designed for bonuses to be equivalent across conditions, in expectation (given optimal spending).

TABLE D.5
ANALYSIS OF POINTS AND BONUSES

	Dependent variable:	
	Points	Bonus
	(1)	(2)
Dining average	−0.772	−0.0003
	p = 0.965	p = 0.980
Low-point dining average	−1.233	−0.003
	p = 0.943	p = 0.789
Budget	−86.393***	−0.077***
	p = 0.00000	p = 0.000
Din. avg. x low din. avg.	10.187	−0.001
	p = 0.552	p = 0.916
Din. avg. x budget	−0.503	0.0004
	p = 0.977	p = 0.971
Low din. avg. x budget	21.489	0.011
	p = 0.210	p = 0.271
Din. x low x budget	−16.603	−0.014
	p = 0.333	p = 0.170
Constant	8,443.077***	0.753***
	p = 0.000	p = 0.000
Observations	821	821
R^2	0.033	0.070
Adjusted R^2	0.025	0.062
Residual Std. Error (df = 813)	489.650	0.286
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Note—Points (left) and bonuses (right) as a function of condition assignment.

There was an effect of being assigned to the budgeting condition, such that budgeters earned fewer points and smaller bonuses. We suspect this relates to the observed tendency towards naïve diversification (the preference to evenly split funds between dining and entertainment budgets). Specifically, if budgeters express a preference for naïve diversification in allocation (allocating near the 50%-50% split), then because budgets predict spending (as previously discussed), then budgeters' spending will be pulled away from the 61%-39% optimal split (corresponding to $140 to dining and $90 to entertainment; see manuscript figure 7 and 8).

Time trends. We designed study 3 as a multiperiod incentivized game in which participants had ample opportunity to learn. One potential concern is whether participants' had not fully learned or developed a decision strategy during the analyzed game weeks. Most concerning would be a pattern reflecting diminishing sensitivity to average value, over time, as participants learned the game dynamics. In exploratory analyses, we do not find evidence of such a pattern. Table D.6 presents five independent analyses of the five incentivized game weeks.

TABLE D.6
WEEKLY REGRESSION RESULTS

	Dependent variable:				
	Dining Share	Dining Share	Dining Share	Dining Share	Dining Share
	(1)	(2)	(3)	(4)	(5)
Dining average	2.319***	2.431***	2.685***	2.121***	3.476***
	p = 0.00005	p = 0.00005	p = 0.00001	p = 0.0003	p = 0.00000
Low-point dining average	0.958+	1.474*	0.537	1.054+	0.139
	p = 0.091	p = 0.013	p = 0.323	p = 0.065	p = 0.820
Din. avg. x low din. avg.	−0.266	0.065	0.312	0.094	0.120
	p = 0.639	p = 0.913	p = 0.566	p = 0.869	p = 0.844
Constant	54.597***	55.118***	54.771***	55.165***	55.265***
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Observations	394	394	394	394	394
R^2	0.049	0.057	0.062	0.043	0.078
Adjusted R^2	0.041	0.050	0.055	0.035	0.070
Residual Std. Error (df = 390)	11.196	11.612	10.749	11.278	12.061
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001				

Note—The dining share of allocation regressed on the contrast-coded condition variables for dining average (high = 1; low = -1), low-point dining average (high = 1; low = -1), and their interactions. Columns 1-5 correspond to incentivized game weeks 1-5. Conceptually, this corresponds to the model to test H1 (table D.1A) with data disaggregated to the weekly level.

Each analysis regresses the dining share of allocation on the dining average and low-point dining average conditions (and their interaction). Across each of these models (corresponding to weeks 1-5), we observe greater allocations to the dining category in the higher dining average condition, and this effect is no smaller in the last week than the first incentivized week.

Building upon this approach, we construct a time-trend outcome variable that reflects the dining share across the five game weeks, each weighted by a linear contrast code to capture time trends. Specifically, we used the contrast weights -2, -1, 0, +1, +2 to correspond to weeks 1, 2, 3, 4, 5, respectively. We considered participants in the budgeting condition, and their dining share in each week was multiplied by the corresponding contrast weights. We then summed this variable. Negative values reflect greater dining allocations *earlier* in the five-week period, and positive values reflect greater dining allocations *later*. We regressed this time-trend on the condition variables (and their interaction). Thus, this approach allows us to test whether the dining share grew or shrank over time based on condition assignment. A positive coefficient on dining average indicates dining allocations grew throughout the five-week game in the high-average condition more than in the low-average condition. A negative coefficient on the low-point dining average indicates dining allocations shrank throughout the five-week game in the high-average condition relative to the low-average condition (in the low-point distribution). There was no tendency for the sensitivity to dining average to decrease across the five incentivized weeks; if anything, there was a marginally significant tendency for sensitivity to dining average to increase.

TABLE D.7
TIME-TREND IN DINING SHARE

	Dependent variable:
	Time-trend outcome
Dining average	0.020^{+}
	p = 0.089
Low-point dining average	-0.021^{+}
	p = 0.080
Din. avg. x low din. avg.	0.008
	p = 0.495
Constant	0.014
	p = 0.239
Observations	394
R^2	0.016
Adjusted R^2	0.009
Residual Std. Error	0.233 (df = 390)
Note:	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Note—The time-trend outcome variable reflects the extent to which a participant's dining share was growing over time (positive value) or shrinking over time (negative value). This was constructed as the sum of the five weekly dining shares, weighted by the contrast weights, corresponding to weeks 1-5: -2, -1, 0, 1, 2. This measure was then regressed on the same set of predictors as in table D.1A for participants in the budgeting condition (used to test H1).