A Model of Relative Thinking: Online Appendices

Benjamin Bushong Michigan State
Matthew Rabin Harvard
Joshua Schwartzstein Harvard Business School

A Further Definitions and Results

A.1 A Basic Result

For any two consumption vectors $c', c \in \mathbb{R}^K$, define $\delta(c', c) \in \mathbb{R}^K$ as a vector that encodes absolute utility differences between c' and c along different consumption dimensions: For all k,

$$\delta_k(c', c) = u_k(c'_k) - u_k(c_k).$$

Choice depends not only on these absolute differences, but also on proportional differences, $\delta_k(c', c)/\Delta_k(C)$. To highlight this, we will consider the impact of "widening" choice sets along particular dimensions. Formally:

Definition A.1. $\tilde{C}$ is *a k-widening of C* if

$$\Delta_k(\tilde{C}) > \Delta_k(C)$$
$$\Delta_i(\tilde{C}) = \Delta_i(C) \text{ for all } i \neq k.$$

In words, $\tilde{C}$ is a k-widening of C if it has a greater range along dimension k and the same range on other dimensions. Although widening may connote set inclusion, our definition does not require this. In our model, the assessment of advantages and disadvantages depends on the range, not on the position within the range or on the position of the range with respect to a reference point.

Proposition A.1. *Let $C, \tilde{C} \subset \mathbb{R}^K$ where $\tilde{C}$ is a k-widening of C.*

1. *Suppose the person is willing to choose c from C. Then for all $\tilde{c} \in \tilde{C}, \tilde{c}' \in \tilde{C}$, and $c' \in C$ such*

that

$$\delta_k(\tilde{c},\tilde{c}') > \delta_k(c,c') > 0$$
$$\frac{\delta_k(\tilde{c},\tilde{c}')}{\Delta_k(\tilde{C})} = \frac{\delta_k(c,c')}{\Delta_k(C)}$$
$$\delta_i(\tilde{c},\tilde{c}') = \delta_i(c,c') \text{ for all } i \neq k,$$

the person will not choose $\tilde{c}'$ from $\tilde{C}$.

2. *Suppose the person is willing to choose c from C. Then for all $\tilde{c} \in \tilde{C}, \tilde{c}' \in \tilde{C}$, and $c' \in C$ such that*

$$\delta_k(c,c') < 0$$
$$\delta_i(\tilde{c},\tilde{c}') = \delta_i(c,c') \text{ for all } i,$$

the person will not choose $\tilde{c}'$ from $\tilde{C}$.

Part 1 of Proposition A.1 says that a person's willingness to choose consumption vector c over consumption vector c' is increasing in the absolute advantages of c relative to c', fixing proportional advantages.[1] But Part 2 says the willingness to choose c is also increasing in its relative advantages, measured in proportion to the range. To illustrate, suppose each c is measured in utility units. Then if the person is willing to choose $c = (2,1,0)$ from $C = \{(2,1,0),(1,2,0)\}$, Part 1 says that he is not willing to choose $\tilde{c}' = (3,2,0)$ from $\tilde{C} = \{(6,1,0),(3,2,0)\}$, which has a bigger range on the first dimension. Part 2 further says that he is not willing to choose $\tilde{c}' = (4,5,3)$ from $\tilde{C} = \{(5,4,3),(4,5,3),(5,0,3)\}$, which has a bigger range on the second dimension.

A.2 Spreading Advantages and Disadvantages

Section 2 supplied examples on how the relative attractiveness of consumption vectors depends on the extent to which their advantages and disadvantages are spread out. To develop formal results, consider the following definition.

Definition A.2. c'' *spreads out the advantages of* c' *relative to* c if there exists a $j \in E(c',c) = \{i : u_i(c'_i) = u_i(c_i)\}$, $k \in A(c',c) = \{i : u_i(c'_i) > u_i(c_i)\}$, and $\varepsilon < \delta_k(c',c)$ such that

$$(u_1(c''_1),\ldots,u_K(c''_K)) = (u_1(c'_1),\ldots,u_K(c'_K)) + \varepsilon \cdot (e_j - e_k),$$

[1] We add the assumption that $\delta_k(\tilde{c},\tilde{c}') > \delta_k(c,c') > 0$ for clarity, but this is implied by $\delta_k(\tilde{c},\tilde{c}')/\Delta_k(\tilde{C}) = \delta_k(c,c')/\Delta_k(C)$ together with $\tilde{C}$ being a k-widening of C.

where e_i is the unit vector whose i'th element is 1. Analagously, c'' *integrates the disadvantages of* c' *relative to* c if there exists $j,k \in D(c',c) = \{i : u_i(c'_i) < u_i(c_i)\}$ such that

$$(u_1(c''_1),\ldots,u_K(c''_K)) = (u_1(c'_1),\ldots,u_K(c'_K)) + \delta_k(c',c)\cdot(e_j - e_k).$$

In words, c'' spreads the advantages of c' relative to c if c'' can be obtained from c' by keeping the total advantages and disadvantages relative to c constant, but spreading its advantages over a greater number of consumption dimensions. Conversely, c'' integrates the disadvantages of c' relative to c if c'' can be obtained from c' by keeping the total advantages and disadvantages relative to c constant, but integrating disadvantages spread over two dimensions into one of those dimensions.

Proposition A.2. *If c'' spreads out the advantages of c' relative to c or integrates the losses of c' relative to c, then $U^N(c'|\{c,c'\}) \geq U^N(c|\{c,c'\}) \Rightarrow U^N(c''|\{c,c''\}) > U^N(c|\{c,c''\})$.*

Proposition A.2 says that, all else equal, the attractiveness of one consumption vector over another goes up when its advantages are spread over more dimensions or its disadvantages are integrated. This connects to the evidence initially derived from diminishing sensitivity of the prospect theory value function that people prefer segregated gains and integrated losses (Thaler 1985), though the evidence on integrated losses (see Thaler 1999) is viewed as far less robust.[2] Thaler gives the following example of a preference for segregated gains: when subjects are asked "Who is happier, someone who wins two lotteries that pay \$50 and \$25 respectively, or someone who wins a single lottery paying \$75?" they tend to believe the person who wins twice is happer. This principle suggests, for example, why sellers of products with multiple dimensions attempt to highlight each dimension separately, e.g., by highlighting the many uses of a product in late-night television advertisements (Thaler 1985).

Turning to losses, Thaler (1985) asked subjects the following question:

> Mr. A received a letter from the IRS saying that he made a minor arithmetical mistake on his tax return and owed \$100. He received a similar letter the same day from his state income tax authority saying he owed \$50. There were no other repercussions

[2]Note that, in contrast to diminishing sensitivity of the prospect theory value function, Proposition A.2 does *not* imply the stronger result that the attractiveness of one consumption vector over another increases in the *degree* to which its advantages are spread or its losses are integrated. For example, while Proposition A.2 implies that if $A = (x,0,0)$ is weakly preferred over $B = (0,0,y)$ from a binary choice set, then $A(\varepsilon) = (x-\varepsilon,\varepsilon,0)$ is strictly preferred over B from a binary choice set, it does not imply that if $A(\varepsilon)$ is weakly preferred over B, then $A(\varepsilon')$ is strictly preferred over B for $0 < \varepsilon < \varepsilon' < x/2$. The intuition is that, in moving from $(x,0,0)$ to $(x-\varepsilon,\varepsilon,0)$, A's advantage of x over B is unambiguously assessed with respect to a lower range: portion $x-\varepsilon$ of the advantage is assessed with respect to range $x-\varepsilon$ rather than x while portion ε is assessed with respect to range ε rather than x. On the other hand, in moving from $A(\varepsilon)$ to $A(\varepsilon')$, there is a trade-off where portion ε of the advantage is now assessed with respect to the *increased* range of ε'. Getting the unambiguous result appears to rely on further assumptions, for example that that $w(\Delta)\cdot\Delta$ is concave in Δ.

from either mistake. Mr. B received a letter from the IRS saying that he made a minor arithmetical mistake on his tax return and owed \$150. There were no other repercussions from his mistake. Who was more upset?

66% of subjects answered "Mr. A", indicating a preference for integrated losses. There is other evidence that urges some caution in how we interpret these results, however. Thaler and Johnson (1990) find that subjects believe Mr. A would be happier if the letters from the IRS and state income tax authority were received two weeks apart rather than on the same day. Under the assumption that events on the same day are easier to integrate, then this pattern goes against a preference for integrated losses. Similarly, while Thaler and Johnson find that subjects say a \$9 loss hurts less when added to a \$250 loss than alone (consistent with a preference for integrating losses), they also say that it hurts *more* when added to a \$30 loss than alone (inconsistent with such a preference). While the overall evidence appears broadly consistent with the predictions of Proposition A.2, the evidence on losses is ambiguous.

The person's sensitivity to the distribution of advantages across dimensions means that, in cases where the dimensions are not obvious, it may be possible to test whether a person treats two potentially distinct dimensions as part of the same or separate consumption dimensions.[3] To adapt an example from Kőszegi and Szeidl (2013, Appendix B), suppose an analyst is uncertain whether a person treats a car radio as part of the same attribute dimension as a car. The analyst can test this question by finding: The price p that makes the person indifferent between buying and not buying the car; the additional price p' that makes the person indifferent between buying the car plus the car radio as opposed to just the car; and finally testing whether the person would buy the car plus the car radio at $p+p'$ dollars. If the person treats the car radio as part of the same attribute dimension as the car, then he will be indifferent. On the other hand, if he treats it as part of a separate dimension then he will not be indifferent; moreover, the direction of preference can identify whether $w(\Delta)$ is decreasing or increasing. Under our model, a person who treats the car radio as a separate attribute will strictly prefer buying the car plus the radio at $p+p'$ dollars because relative thinking implies that the person prefers segregated advantages.[4]

[3] This feature of the model also means that choice behavior can exhibit cycles, suggesting caution in how we elicit preferences from relative thinkers. To take an example, suppose consumption is expressed in utils and consider $c=(2,2,0)$, $c'=(4,0,0)$, and $c^y=(0,0,y)$. A relative thinker expresses indifference between c and c' from a binary choice set (because this is a "balanced choice"), as well as between c' and $c^{y=4}$. But, because the person acts as if he prefers segregated advantages, he expresses a strict preference for c over $c^{y=4}$ from a binary choice set. Interpreting the third dimension as "money" and the first two dimensions as attributes, the person would express indifference between the attribute components of c and c' if asked to choose between them but appear to place a greater dollar value on c. Under the view that $U(\cdot)$ is the correct welfare metric, the binary choice elicitation yields the more accurate conclusion in this example.

[4] In Kőszegi and Szeidl's (2013) model, such a person will strictly prefer not to buy at this price because of their bias towards concentration. Likewise, diminishing sensitivity does not share this prediction with a reference point of zero, but rather implies that the person would be indifferent.

The model more broadly implies that it is easier to advantageously frame items whose advantages are more spread out:

Proposition A.3. *Assume that $u_k(\cdot)$ is unbounded below for each k. Let $c, c', c'' \in \mathbb{R}^K$ where c'' spreads out the advantages of c' relative to c. Supposing there is a C containing $\{c, c'\}$ such that c' is chosen from C, then there is a $\tilde{C}$ containing $\{c, c''\}$ such that c'' is chosen from $\tilde{C}$.*

A.3 Further Results on the Limits of Choice-Set Effects and Prophylactic Decoys

Examining the necessary and sufficient condition (1) yields the following corollary:

Corollary A.1.

1. *If c dominates c', where $c, c' \in \mathbb{R}^K$, then there does not exist a C containing $\{c, c'\}$ such that c' would be chosen from C.*
2. *Consider $c, c' \in \mathbb{R}^K$ where the total advantages of c' relative to c satisfy $\delta_A(c', c) \equiv \sum_{i \in A(c',c)} \delta_i(c', c) = \tilde{\delta}_A$ for some $\tilde{\delta}_A > 0$. Then, additionally assuming* N3*, there exists a finite constant $\bar{\delta} > 0$ for which there is a C containing $\{c, c'\}$ such that c' is chosen from C only if the total disadvantages of c' relative to c satisfy $\delta_D(c', c) \equiv -\sum_{i \in D(c',c)} \delta_i(c', c) < \bar{\delta}$.*

The first part of the corollary says that dominated options can never be framed in a way where they will be chosen over dominating alternatives. The second says that it is only possible to frame an inferior option in a way that it is chosen over a superior alternative if its disadvantages are not too large relative to its advantages.

The previous result establishes one way in which the impact of the comparison set is bounded in our model. The next result establishes another: for any option c, there exists a choice set containing c such that c will be chosen and, for any expansion of that set, only options that yield "roughly equivalent" utility to c or better can be chosen. Recalling that $\delta_A(c', c) = \sum_{i \in A(c',c)} \delta_i(c', c)$ and $\delta_D(c', c) = -\sum_{i \in D(c',c)} \delta_i(c', c)$, we have the following result:

Proposition A.4. *Assume N3 and that $u_k(\cdot)$ is unbounded below for each k. For any $c \in \mathbb{R}^K$ and $\varepsilon > 0$, there exists some C_ε containing c such that the person would be willing to choose c from C_ε and would not choose any $c' \in \mathbb{R}^K$ with $\delta_A(c', c) = 0$ or $\delta_A(c', c) > 0$ and*

$$\frac{\delta_D(c', c)}{\delta_A(c', c)} - 1 > \varepsilon$$

from any $\tilde{C}$ containing C_ε.

Proposition A.4 says that, for any option c, it is possible to construct a choice set containing c as well as "prophylactic decoys" that would not be chosen, but prevent expanding the choice set in ways that allow sufficiently inferior options to c to be framed as being better. With unbounded utility, it is always possible to add options that make the ranges on dimensions sufficiently large such that further expanding the choice set will not make some dimensions receive much larger decision weights than others. For example, if $c = (1,8,2)$ and $\bar{u} > 0$, then c is chosen from $C = \{(1,8,2),(1,8-\bar{u},2-\bar{u}),(1-\bar{u},8,2-\bar{u}),(1-\bar{u},8-\bar{u},2)\}$ and, as $\bar{u} \to \infty$, it is impossible to expand C in a way that significantly alters the ranges along various dimensions and allows an inferior option to c to be chosen.

These ideas may be seen more clearly when we start from two options rather than one. A simple corollary is that when one option c has a higher un-normed utility than another c', it is possible to find a comparison set including those options such that the person chooses c from that set and where it is not possible to expand the set in a way that will reverse his preference.

Corollary A.2. *Assume the conditions of Proposition A.4 hold. For any $c, c' \in \mathbb{R}^K$ with $U(c) > U(c')$, there exists some C containing $\{c, c'\}$ such that the person would be willing to choose c from C and would not choose c' from any $\tilde{C}$ containing C.*

Again, applying the result to think about product market competition, this result says that if a firm has a superior product to a competitor then, with unbounded utility, it can always add inferior decoys that lead the consumer to choose its target product, and prevent the competitor from adding decoys that frame its inferior product as superior.[5] To take an example, consider $c = (8,2)$ and $c' = (4,7)$. For concreteness, we could imagine cars where c has better speed and c' has better comfort. Starting from a binary choice set, the speedy car producer may be able to get consumers to buy its inferior product by adding similarly speedy but really uncomfortable decoy cars. However, Corollary A.2 tells us that the comfortable car producer can always add prophylactic decoy cars that prevent the speedy car producer from being able to do this. These prophylactic decoys, such as $(-\bar{u}, 7.1)$ for $\bar{u}$ large, would "double-down" on the comfortable car's speed disadvantage, protecting this disadvantage from being framed as all that bad.[6]

[5] The key assumption is that the superior firm can add decoys that make the range on its disadvantageous dimensions sufficiently large that the inferior firm cannot add its own decoys that significantly magnify the relative weight placed on its advantageous dimensions. This can be satisfied with bounded utility as well, so long as lower bounds of utilities along the superior firm's advantageous dimensions weakly exceed lower bounds along its disadvantageous dimensions. If we were to relax assumption *N3* that $w(\infty) > 0$ then, with unbounded utility, we could instead observe a form of "instability" where it is possible to expand any set C from which c is chosen so that c' is chosen and vice-versa.

[6] We suspect a similar result also holds for Kőszegi and Szeidl's (2013) model under natural restrictions on the "focusing weights", though the prophylactic decoys would look different.

B Eliciting Model Ingredients from Behavior

This section outlines an algorithm for eliciting $u_k(\cdot)$ and $w(\cdot)$ from behavior. The elicitation essentially follows the steps laid out by Kőszegi and Szeidl (2013) to elicit the ingredients of their model; we will closely follow their presentation. Their algorithm works for us because our model shares the feature that people make consumption-utility-maximizing choices in "balanced" decisions, which allows us to elicit consumption utility by examining choices in such decisions. We then elicit the weighting function $w(\cdot)$ by examining how bigger ranges influence the person's sensitivity to given differences in consumption utility.

We assume *N0(d)* and *N2* (but do not impose *N1*) and follow Kőszegi and Szeidl (2013) by assuming that we know how options map into attributes, that we can separately manipulate individual attributes of a person's options, and that the utility functions $u_k(\cdot)$ are differentiable. We also, without loss of generality, normalize $u_k(0) = 0$ for all k, $u_1'(0) = 1$, and $w(1) = 1$. We depart from Kőszegi and Szeidl (2013) by assuming $w(\Delta) \cdot \Delta$ is strictly increasing (Assumption *N2*), while they make the stronger assumption that $w(\Delta)$—or $g(\Delta)$ in their notation—is strictly increasing. We will see that their elicitation algorithm still works under our weaker assumption and, in fact, their elicitation can be used to test our assumption that $w(\Delta)$ is decreasing against theirs that $w(\Delta)$ is increasing.

The first step of the algorithm is to elicit the utility functions $u_k(\cdot)$. Restricting attention to dimensions 1 and k, consider choice sets of the form

$$\{(0, x+q), (p, x)\}$$

for any $x \in \mathbb{R}$ and $p > 0$. For $p > 0$, set $q = q_x(p)$ to equal the amount that makes a person indifferent between the two options, so

$$w(u_1(p) - u_1(0)) \cdot (u_1(p) - u_1(0)) = w(u_k(x + q_x(p)) - u_k(x)) \cdot (u_k(x + q_x(p)) - u_k(x)),$$

which because $w(\Delta) \cdot \Delta$ is strictly increasing in Δ, implies that

$$u_1(p) - u_1(0) = u_k(x + q_x(p)) - u_k(x).$$

Dividing by p and letting $p \to 0$ yields

$$u_1'(0) = u_k'(x) \cdot q_x'(0),$$

which (using the normalization that $u_1'(0) = 1$) gives $u_k'(x)$, and (using the normalization that $u_k(0) = 0$) gives the entire utility function $u_k(\cdot)$. Intuitively, this step of the algorithm gives us, for every x, the marginal rate of substitution of attribute 1 for attribute k at $(0,x)$—this is $q'_x(0) = u'_1(0)/u'_k(x)$—which yields the entire shape of $u_k(x)$ given the normalization that $u'_1(0) = 1$. We can then similarly recover $u_1(\cdot)$ through using the elicited utility function for some $k > 1$.

The second step of the algorithm elicits the weights $w(\cdot)$, where we can now work directly with utilities since they have been elicited. Focus on dimensions $1,2$, and 3, and consider choice sets of the form

$$\{(0,0,x_0),(1,x-p,0),(1-q,x,0)\},$$

for any $x \in \mathbb{R}^+$, where $x_0 > 0$ is sufficiently low that $(0,0,x_0)$ will not be chosen and whose purpose is to keep this option from being dominated by the others and from lying outside the comparison set (this is the only step of the algorithm where having more than two attributes matters). For some $p \in (0,x)$, we now find the $q = q_x(p)$ that makes the person indifferent between the second two options in the choice set, requiring that p is sufficiently small that $q_x(p) < 1$, so

$$w(1)\cdot 1 + w(x)\cdot(x-p) = w(1)\cdot(1-q_x(p)) + w(x)\cdot x.$$

This implies that $w(x)\cdot p = w(1)\cdot q_x(p)$ and, by the normalization $w(1) = 1$, gives us

$$w(x) = \frac{q_x(p)}{p}.$$

In this manner, we can elicit the entire weighting function $w(\cdot)$. Intuitively, for all x, this step of the algorithm elicits the marginal rate of substitution of utils along a dimension with weight $w(x)$ for utils along a dimension with weight $w(1)$, which yields exactly $w(x)$ given the normalization $w(1) = 1$.

With this elicited weighting function, we can, for example, test our assumption that $w(\cdot)$ is decreasing against Kőszegi and Szeidl's (2013) that $w(\cdot)$ is increasing. To illustrate, suppose dimensions 1 and 2 represent utility as a function of the number of apples and oranges, respectively, where utility is elicited through the first step of the algorithm. Ignoring the third dimension for simplicity, if we see that the person strictly prefers $(1/2 \text{ utils apples}, 3 \text{ utils oranges})$ from the choice set

$$\{(0 \text{ utils apples}, 0 \text{ utils oranges}), (1 \text{ utils apples}, 2.5 \text{ utils oranges}), (1/2 \text{ utils apples}, 3 \text{ utils oranges})\},$$

then $w(3)/w(1) > 1$, consistent with Kőszegi and Szeidl (2013), while if the person instead strictly prefers $(1 \text{ utils apples}, 2.5 \text{ utils oranges})$ from this choice set, then instead $w(3)/w(1) < 1$, consistent with our model.

C More Detailed Comparison to Other Models

As noted in the introduction, the basic feature of our model —that a given absolute difference looms smaller in the context of bigger ranges (Parducci 1965)—is not shared by Bordalo, Gennaioli and Shleifer (2012, 2013) or other recent approaches by Kőszegi and Szeidl (2013) and Cunningham (2013), who model how different features of the choice context influence how attributes of different options are weighed. To enable a detailed comparison between the approaches, we present versions of their models using similar notation to ours, and compare the models in the context of simple examples along the lines of the one introduced in Section 3.

All of these models share the feature that there is some $U(c) = \sum_k u_k(c_k)$ that is a person's consumption utility for a K-dimensional consumption bundle c, while there is some $\hat{U}(c|C) = \sum_k w_k \cdot u_k(c_k)$ that is the "decision consumption utility" that he acts on. The models by Bordalo, Gennaioli and Shleifer (2012, 2013), Kőszegi and Szeidl (2013), and Cunningham (2013) differ from each other's, and from ours, in how they endogeneize the "decision weights" $w_k > 0$ as functions of various features of the choice context, and possibly the option c under consideration. Specifically, their models assume the following:

Alternative Model 1 (Bordalo, Gennaioli and Shleifer (2012, 2013).)**.** Bordalo, Gennaioli, and Shleifer's (2013) model of salience in consumer choice says that for option c, $w_i > w_j$ if and only if attribute i is "more salient" than j for option c given "evoked" set C of size N, where "more salient" is defined in the following way. Ignoring ties and, for notational simplicity, assuming positive attributes (each $u_k(c_k) > 0$), attribute i is more salient than j for c if $\sigma_i(c|C) > \sigma_j(c|C)$, where $\sigma_k(c|C) \equiv \sigma\left(u_k(c_k), \frac{1}{N}\sum_{c' \in C} u_k(c'_k)\right)$ and $\sigma(\cdot,\cdot)$, the "salience function", is symmetric, continuous, and satisfies the following conditions (thinking of $\bar{u}_k = \bar{u}_k(C) \equiv \frac{1}{N}\sum_{c' \in C} u_k(c'_k)$):

1. *Ordering.* Let $\mu = \text{sign}(u_k - \bar{u}_k)$. Then for any $\varepsilon, \varepsilon' \geq 0$ with $\varepsilon + \varepsilon' > 0$,

$$\sigma(u_k + \mu\varepsilon, \bar{u}_k - \mu\varepsilon') > \sigma(u_k, \bar{u}_k).$$

2. *Diminishing Sensitivity.* For any $u_k, \bar{u}_k \geq 0$ and for all $\varepsilon > 0$,

$$\sigma(u_k + \varepsilon, \bar{u}_k + \varepsilon) \leq \sigma(u_k, \bar{u}_k).$$

3. *Homogeneity of Degree Zero.* For all $\alpha > 0$,

$$\sigma(\alpha \cdot u_k, \alpha \cdot \bar{u}_k) = \sigma(u_k, \bar{u}_k).$$

The ordering property implies that, fixing the average level of an attribute, salience is increasing in absolute distance from the average. The diminishing-sensitivity property implies that, fixing the absolute distance from the average, salience is decreasing in the level of the average. Note that these two properties can point in opposite directions: increasing $u_i(c)$ for the option c with the highest value of $u_i(\cdot)$ increases $(u_i(c) - \bar{u}_i)$, suggesting higher salience by ordering, but also increases $\bar{u}_i$, which suggests lower salience by diminishing sensitivity. Homogeneity of degree zero places some structure on the trade-off between these two properties by, in this example, implying that ordering dominates diminishing sensitivity if and only if $u_i(c)/\bar{u}_i$ increases.

More generally, using Assumptions 1-3, it is straightforward to show the following:[7]

$$\text{For option } c,\ w_i > w_j \Leftrightarrow \sigma_i(c|C) > \sigma_j(c|C) \Leftrightarrow \frac{\max\{u_i(c_i), \bar{u}_i(C)\}}{\min\{u_i(c_i), \bar{u}_i(C)\}} > \frac{\max\{u_j(c_j), \bar{u}_j(C)\}}{\min\{u_j(c_j), \bar{u}_j(C)\}}, \tag{BGS}$$

where the level of w_i depends only on the salience rank of attribute i for option c in comparison set C.[8] An interpretation of condition (BGS) is that attribute i of option c attracts more attention than attribute j and receives greater "decision weight" when it "stands out" more relative to the average level of the attribute, where it stands out more when it is further from the average level of the attribute in proportional terms.

Alternative Model 2 (Kőszegi and Szeidl (2013))**.** Kőszegi and Szeidl's (2013) model of focusing specifies that the decision weight on attribute k equals

$$w_k = g(\Delta_k(C)),\ g'(\cdot) > 0, \tag{KS}$$

where $\Delta_k(C) = \max_{c' \in C} u_k(c'_k) - \min_{c' \in C} u_k(c'_k)$ equals the range of consumption utility along dimension k, exactly as in our model. However, the weight on a dimension is assumed to be *increasing* in this range, $g'(\Delta) > 0$, which directly opposes Assumption *N1* of our model. An interpretation of condition (KS) is that people focus more on attributes in which options generate a "greater range" of consumption utility, leading people to attend more to fixed differences in the context of bigger ranges.

Alternative Model 3 (Cunningham (2013))**.** Cunningham (2013) presents a model of relative

[7] As Bordalo, Gennaioli and Shleifer (2013) discuss, Assumption 2 (Diminishing Sensitivity) is actually redundant given Assumptions 1 and 3 (Ordering and Homogeneity of Degree Zero).

[8] Specifically, letting $r_i(c|C) \in \{1, \ldots, K\}$ represent the salience rank of attribute i for option c given comparison set C (the most salient attribute has rank 1), Bordalo, Gennaioli and Shleifer (2013, Appendix B) assume that the weight attached to attribute i for option c is given by

$$w_i = \frac{\delta^{r_i(c|C)}}{\sum_k \delta^{r_k(c|C)}},$$

where $\delta \in (0, 1]$ inversely parameterizes the degree to which the salience ranking matters for choices.

thinking in which a person is less sensitive to changes on an attribute dimension when he has encountered larger absolute values along that dimension. Cunningham's model is one in which previous choice sets, in addition to the current choice set, affect a person's decision preferences, so we need to make some assumptions to compare the predictions of his model to ours, and in particular to apply his model when a person's choice history is unknown. We will apply his model assuming that the person's choice history equals his current choice or comparison set C. It is then in the spirit of his assumptions that the decision weight attached to attribute k equals[9]

$$w_k = f_k(|\bar{u}_k(C)|),\ f_k'(\cdot) < 0\ \forall k, \tag{TC}$$

where $\bar{u}_k(C) = \frac{1}{N}\sum_{c' \in C} u_k(c'_k)$ is the average value of attribute k across elements of C and N is the number of elements in C. Formulation (TC) says that a person is less sensitive to differences on an attribute dimension in the context of choice sets containing options that, on average, have larger absolute values along that dimension.

To illustrate differences between the models, return to the example introduced in Section 3. Suppose a person is deciding between the following jobs:

Job X. Salary: 100K, Days Off: 199
Job Y. Salary: 110K, Days Off: 189
Job Z. Salary: 120K, Days Off: 119,

where his underlying utility is represented by $U = \text{Salary} + 1000 \times \text{Days Off}$. First, we will consider the person's choice of jobs when he is just choosing between X and Y, and then we will consider his choice when he can also choose Z.

As noted in Section 3, our model predicts that the person will be indifferent between jobs X and Y when choosing from $\{X,Y\}$, but instead strictly prefers the higher salary job Y when choosing from $\{X,Y,Z\}$. None of the three other models share our prediction in this example. The predictions of Kőszegi and Szeidl's (2013) model were presented in Section 3. Bordalo, Gennaioli, and Shleifer's (2013) predicts that a person will strictly prefer choosing the higher salary job Y from $\{X,Y\}$: Using condition (BGS), we see that salary is more salient than days off for both options in $\{X,Y\}$—salary is more salient than days off for X since $105/100 > 199/194$, and salary is more salient than days off for Y since $110/105 > 194/189$—so the person places greater decision weight on salary

[9] Cunningham (2013) considers a more general framework where utility is not necessarily separable across dimensions, and makes assumptions directly on marginal rates of substitution. Part of his paper considers implications of weaker assumptions on how "translations" of histories along dimensions influence marginal rates of substitution, rather than average levels of attributes along dimensions. We focus on his average formulation because it enables sharper predictions across a wider range of situations: it is always possible to rank averages, but not always possible to rank histories by translation.

and chooses the higher salary option. Intuitively, by diminishing sensitivity, a 5K utility difference relative to the average on the salary dimension stands out more than a 5K utility difference relative to the average on the days off dimension, as the average on the salary dimension is lower. Like Kőszegi and Szeidl (2013), Bordalo, Gennaioli, and Shleifer (2013) predict that the person will reverse her choice to X from $\{X,Y,Z\}$: Using condition (BGS), the addition of Z leads days off to be salient for all options—days off is more salient than salary for X since $199/169 > 110/100$, days off is more salient than salary for Y since $189/169 > 110/110$, and days off is more salient than salary for Z since $169/119 > 120/110$—so the person places greater decision weight on days off and chooses X. Intuitively, their model says that the addition of job Z, which is a relative outlier in terms of days off, causes the days off of the various options to really stand out. Like Kőszegi and Szeidl (2013), the salience-based prediction of Bordalo, Gennaioli, and Shleifer (2013) in this two-dimensional example seems at odds with intuition generated from laboratory evidence on attraction or range effects.[10]

Cunningham's (2013) formulation does not pin down what a person chooses from $\{X,Y\}$ (since the function governing the decision weights can vary across k), but says that if the person is initially indifferent between X and Y, then the addition of Z would lead him to choose X: Since the addition of Z brings up the average on the salary dimension and brings down the average on the days off dimension, condition (TC) tells us that it leads the person to care less about salary relative to days off, thereby making X look more attractive than Y. Cunningham's average-based formulation yields opposite predictions to our range-based formulation when, like in this example, adding an option impacts averages and ranges in different directions.

We can re-frame this example slightly to illustrate another point of comparison. Suppose that a person frames the jobs in terms of salary and vacation days, rather than salary and days off, where vacation days equal days off minus weekend days (with roughly 104 weekend days in a year). The idea is that the person's point of reference might be to be able to take off all weekend days rather than to take off no days. Then the problem can be re-written as choosing between the following jobs:

Job X. Salary: 100K, Vacation Days: 95
Job Y. Salary: 110K, Vacation Days: 85
Job Z. Salary: 120K, Vacation Days: 15,

[10]As Bordalo, Gennaioli and Shleifer (2013) note, their model accommodates the attraction effect when people choose between options that vary in quality and price. For example, people will choose $(70,-20)$ from $\{(70,-20),(80,-30)\}$ but will instead choose $(80,-30)$ from $\{(70,-20),(80,-30),(80,-40)\}$ since the addition of the decoy option $(80,-40)$ makes the price of the middle option not salient because it equals the average price. However, their model also robustly accommodates the opposite effect — call it the "repulsion effect" — in these situations. In particular, the person will choose $(70,-20)$ if the price of decoy is made larger so price becomes salient for the middle option. For example, the person will choose $(70,-20)$ from $\{(70,-20),(80,-30),(80,-70)\}$.

where the person's underlying utility is represented by $U = \text{Salary} + 1000 \times \text{Vacation Days}$. This change in formulation does not influence the predictions of our model, or of Kőszegi and Szeidl's (2013), on how the person chooses from $\{X, Y\}$ or from $\{X, Y, Z\}$ because this change does not affect utility ranges along the different dimensions. On the other hand, this change does influence the predictions of Bordalo, Gennaioli, and Shleifer (2013). Specifically, it alters the prediction of which choice the person makes from $\{X, Y\}$ because diminishing sensitivity is defined relative to a reference point: With the new reference point, vacation days are now more rather than less salient for both options in $\{X, Y\}$ because a 5K difference looms larger relative to an average of 90K than 105K, implying that a person chooses X rather than Y from the binary choice set.[11] And while this particular change in the reference point does not alter the qualitative predictions of Cunningham (2013), a different change does: Suppose a person uses a reference point where all 365 days are taken off and each option is represented in terms of (Salary, Work Days), where utility is represented by $U = \text{Salary} - 1000 \times \text{Work Days}$. In this case, Cunningham (2013) says that the addition of Z reduces the person's sensitivity to work days, since it raises the average number of such days, while the earlier framing in terms of days off instead suggested that the addition of Z would increase the person's sensitivity to work days since it decreased the average number of days off.

D A Sketch of How Focusing and Relative Thinking Could Be Combined

For a formulation that combines focusing and relative thinking, order the dimensions $k = 1, 2, ...K$ (where K can be either finite or infinite) such that $\Delta_k \geq \Delta_{k+1}$. Then, choosing parameter $\chi \in [0, 1]$, let ϕ_k be the *approximate focus weight* on dimension k as given by $\phi_{k+1} = \chi \cdot \phi_k$ for $k < K$, $\phi_K = \phi_{K-1}$, and $\sum_{k=1}^{K} \phi_k = 1$. The actual focus weights would be modified to take into account exact ties, where $\Delta_k = \Delta_{k+1}$. We denote the true focus weights by g_k such that $\sum_{k=1}^{j} g_k = \sum_{k=1}^{j} \phi_k$ for all j where $\Delta_j > \Delta_{j+1}$, and $g_j = g_{j+1}$ where $\Delta_j = \Delta_{j+1}$. Finally, we replace the weighting functions w_k, previously given by $w_k = w(\Delta_k(C))$. Instead, they are given by $w_k = g_k \cdot w(\Delta_k(C))$, where $w(\cdot)$ follows our Norming Assumptions *N0-N3*.

This implies (by brute-force construction) that range effects in two dimensional settings will be determined solely by relative thinking. But if there are at least three dimensions, the focus weights can matter. If we assume $\chi = .5$, for instance, three dimensions with utility ranges $(3, 2, 1)$ would get focus weights $(g_1, g_2, g_3) = (\frac{4}{8}, \frac{2}{8}, \frac{2}{8})$; dimensions with utility ranges $(3, 3, 1)$ would get

[11] Specifically, given $C = \{X, Y\}$, vacation days are salient for X since $95/90 > 105/100$ and are salient for Y since $90/85 > 110/105$.

focus weights $(g_1, g_2, g_3) = (\frac{3}{8}, \frac{3}{8}, \frac{2}{8})$; and dimensions with utility ranges $(3, 1, 1, 1, 1)$ would get focus weights $(g_1, g_2, g_3, g_4, g_5) = (\frac{4}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$. Comparing the first two examples illustrates that focus weights increase in the range when the increase influences the ranking of ranges across dimensions; comparing the second and third illustrates that this formulation shares Kőszegi and Szeidl's (2013) key feature that people pay less attention to advantages which are more spread out. In this (admittedly crude) formulation, relative thinking will dominate in two-dimensional choices or whenever an increase in the range does not influence the ranking of ranges across dimensions. While big differences can draw attention (bigger Δ_k increases g_k), range-based relative thinking will dominate conditional on the allocation of attention (bigger Δ_k decreases w_k, conditional on g_k).

E Experiments: Pilots and Instructions

In this section, we provide the full text of our experimental instructions. We first present two pilots that were used to assess whether our vignettes and real-effort task were well-calibrated. Throughout this section, we use parentheses to denote alternative instructions corresponding to different treatments. Brackets denote experimental logic not visible to the participant but presented here to aid the reader. All experiments commenced with an informed consent form. This research was approved by the Michigan State and Harvard IRBs.

E.1 Pilot Experiments

We conducted two pilot experiments that were intended to help us calibrate the experimental designs without testing our hypotheses.

Our first pilot followed the instructions of Experiment 1 verbatim until reaching the vignettes. The vignettes were replaced with the following alternative vignettes, which appeared in randomized order:

> Vignette 1. You went to a store to buy a phone, expecting to pay \$500 and expecting the trip to the store to last 30 minutes. You've spent 30 minutes at the store and it turns out the phone you want costs exactly \$500. Right before buying it, the clerk at the check-out counter says you could save \$10 on the phone by filling out a 15-minute survey, bringing the total time at the store up to 45 minutes and the total cost of the phone down to \$490. Do you fill out the survey?
>
> Vignette 2. Tarso has been working at the same restaurant for years and after a long shift, he is ready to go home. As he packs his things, he notices a group of four

people enter and sit at a booth in the corner. He quickly thought to himself: should I stay or should I go? His work shares tips and he suspects that he would earn a $5 tip from the remaining table. Tarso is uncertain about his tips thus far, but suspects he earned about $50 thus far from the pooled tips tonight. Therefore if he stays, Tarso thinks he'll earn about $55. Do you think he [stayed and served one final table] [left for the night]

Vignette 3. Imagine you have spent the day shopping. One item you have been shopping for is a pair of headphones. At the end of the day, you find yourself at a store that has the brand and model you want for $100. This is a good price but not the best you have seen today. One store—a thirty minute detour on your way home—has them for $75. Do you buy the $100 headphones and go home, or do you instead decide to take the detour to buy them for $75 at the other store?

Vignette 4. Suppose that you are a new personnel manager in a company. Your duties include setting the hourly wages for current employees based on their "work quality index" (0 - 100, the higher the better) and number of sick days in the last year (the fewer, the better). On your first day, two employees comes into your office to request a raise. When you start to examine the company records you find the old personnel manager left them completely disorganized. You only have the following data for each employee:
Employee's Name: J.M.
Work Quality Index: 62
Number of Sick Days: 20
Employee's Name: K.B.
Work Quality Index: 48
Number of Sick Days: 8

You tell each employee that you will have a decision by the end of the week, knowing that you can offer only two raises. Your superiors realize that you are working with limited information. Still, they expect you to apply your best judgment and decide which employee deserves to be offered a raise. Knowing only the information provided above, which employee deserves a raise?

Vignette 5. Kevin and Jenny are employees at a major electronics store. Both have been considering purchasing new TVs; because of their seniority, each faces a different possible discount.
Kevin can get a 15% TV discount and is considering a $1500 TV.

Jenny can get 25% TV discount and is considering a $900 TV.

Both Kevin and Jenny have given this some thought. Despite references to Kevin and Jenny above, please ignore the question that follows and choose [Do not choose this answer]. Do not choose either Kevin or Jenny. This is an attention check. Who do you think was more likely to purchase the TV: Kevin or Jenny?

Discussion of first pilot. We briefly present results of the vignette pilot. The first vignette was designed to ensure that a reasonable number of people would say they were willing to fill out the survey (while not directly testing our theory-guided hypotheses). 38 out of 90 participants indicated they were willing to do so. The second vignette was intended to ensure that $5 was enough to continue to work – 46 out of 90 participants indicated they thought Tarso would stay and serve another group. The third vignette was a modernization of the jacket-calculator problem which was useful to check whether subjects were behaving as expected. 72 out of 89 participants indicated they would drive to the other store (consistent with our priors). One participant chose the alternative "Do not choose this option" (which was present for each vignette) and thus failed the attention check on this question. The fourth question was an early attempt at a two-dimensional design, where we intended J.M. and K.B. to be equally attractive for the raise. However, 56 out of 89 participants chose J.M., and we were unsure how to proceed with the design and therefore did not include this sort of question in our final experiment. The last vignette was an attention check – 85 out of 90 participants correctly chose "Do not choose this answer".[12]

Our second pilot followed the instructions from Experiment 2 exactly until the list of questions; participants then faced the following menu of options, which we label with letters here for ease of reference:

$$\begin{aligned} A &: (\$2.00, 12 \text{ tasks}) \\ B &: (\$3.00, 20 \text{ tasks}) \\ C &: (\$2.20, 14 \text{ tasks}) \\ D &: (\$2.55, 17 \text{ tasks}). \end{aligned}$$

Discussion of second pilot. The specific questions above were designed to identify whether participants were roughly indifferent to A and B. Notably, options C and D do not extend the range on either money or effort when added to a choice set containing A and B. Moreover, the pilot was meant to check whether participants would understand the experimental instructions (e.g., to rank

[12]Out of the five participants who incorrectly chose a real alternative in this attention check question, two indicated in the free response question that they found the wording confusing. A third emailed the authors and similarly indicated confusion. As a result, we did not include this question in the final experiment, but maintained the question-by-question attention checks with "Do not choose this option."

options), and to gauge the difficulty of the real-effort task. It included an open-ended question about participant confusion. ("Were you confused by any questions? If so, feel free to explain below.")

We set an ex-ante heuristic that so long as a "roughly equal" number of participants chose A and B as their most preferred alternative, we could proceed. Out of 115 participants, 67 indicated that alternative A was their most preferred; 45 indicated that alternative B was their most preferred; 2 indicated that alternative C was their most preferred, and 1 indicated that alternative D was their most preferred. Taking this feedback, we decided to "move" A and B slightly closer together in monetary amounts in our final design. (In our write-up of the final design, the revised versions of A and B are re-labeled B and C, respectively.)

E.2 Complete Experimental Instructions: Experiment 1

In this survey, we will ask you a series of hypothetical questions. Each question describes a real-life scenario and we want to know what you would do in that scenario (or what you think someone else would do).

There are a total of three hypothetical questions, so please take your time to read each closely. In order to ensure you read the options closely, we have added options that say "Do not choose this option". If you choose any of those options, you will forfeit your payment for participation.

Additionally, a timer will prevent responses before 15 seconds have elapsed to ensure that you review the options carefully.

At the end of the hypothetical questions, we will ask a few short demographic questions and the survey will be complete.

[Order of the following three vignettes was randomized]

- Imagine you have spent the day shopping. One item you have been shopping for is a laptop (pair of headphones). At the end of the day, you find yourself at a store that has the brand and model you want for $1000 ($100). This is a good price but not the best you have seen today. One store—a thirty minute detour on your way home—has them for $975 ($75). Do you buy the $1000 laptop ($100 headphones) and go home, or do you instead decide to take the detour to buy it (them) for $975 ($75) at the other store?

  [Order of the response buttons was randomized]

  ☐ Buy it for $1000 (them for $100) and go home

  ☐ Drive 30 minutes and buy it for $975 (them for $75)

  ☐ Do not choose this option

- You went to a store to buy a phone, expecting to pay between $490 ($190) and $510 ($810) and expecting the trip to the store to last between 5 (15) and 55 (45) minutes. You've spent 30 minutes at the store and it turns out the phone you want costs exactly $500. Right before buying it, the clerk at the check-out counter says you could save $5 on the phone by filling out a 15-minute survey, bringing the total time at the store up to 45 minutes and the total cost of the phone down to $495. Do you fill out the survey?

☐ Yes

☐ No

☐ Do not choose this option

- Tarso has been working at the same restaurant for years and after a long shift, he is ready to go home. As he packs his things, he notices a group of four people enter and sit at a booth in the corner. He quickly thought to himself: should I stay or should I go? His work shares tips and he suspects that he would earn a $5 tip from the remaining table. Tarso is uncertain about his tips thus far, but suspects he earned between $35 ($15) and $40 ($60) thus far from the pooled tips tonight. Therefore if he stays, Tarso thinks he'll earn between $40 and $45 ($20 and $65). Do you think he:

  ☐ Stayed and served one final table

  ☐ Left for the night

  ☐ Do not choose this option

We will conclude with some simple demographic questions.

What is your gender?

☐ Male
☐ Female

What is your annual income?

☐ less than $15,000
☐ $15,000 - $29,999
☐ $30,000 - $59,999
☐ $60,000 - $99,999
☐ $100,000 or more

What is your age (in years)?

What is your zip code? [Format: 00000]

You have reached the end of the survey. Please advance to the final screen to submit your work. Note that the MTurk code you need is on the screen that follows.

You will be paid within one week of completing the study.

E.3 Complete Experiment Instructions: Experiment 2

We will not deceive you whatsoever in this experiment. All of the instructions provide examples and guidance for the actual tasks you will do. There will be no tricks.

You will do a simple job that takes roughly 5 minutes. You will earn a fixed payment of $1 for completing this job and reading the instructions carefully.

We will then ask you some questions about your willingness to do additional work for additional pay.

You must complete the session to earn any pay for this study. There will be absolutely no exceptions to this rule. All payments will be credited to your MTurk account within one week of completing the study.

Overview

1. We will review the real-effort task and you will complete approximately five minutes of work. 2. We will ask you about doing additional work for additional pay or to avoid reducing your payment. 3. If you elect to do additional work for additional pay, you will complete that additional work and the survey will be complete.

You will know that you have reached the end of the survey when you see a screen saying "THIS IS THE END OF THE SURVEY". Please do not exit until you have seen this screen. This final screen includes a code that you must input into MTurk in order to get paid.

Task

Your task is to count. You will see an image like the image below:

!	0	)	0	0	0	?	(	i	!	0	)	!!	!	0
!!	0	t	!	0	t	!	)	t	0	(	!	t	(	?
!	!	!	t	)	!	!	(	!	0	!	!	0	!	!!
!!	(	!	)	!	!	t	t	?	!	!	0	0	!	1
!	!	0	0	!	(	)	0	t	!	)	?	0	0	i
t	1	!	!	0	!	0	!!	!	t	!	!	!	!	0
t	!	(	!	!	0	!	(	(	!	!!	!	!	t	?
!	0	!	0	0	0	!	0	!	!	)	!	i	!	0
(	!	!	!	!	t	!	0	!	!	!	!	!	(	!
0	(	0	(	0	)	!	!	i	0	(	t	t	0	0

You will then be asked to count a specific character that is present in the image. The question will be phrased as:
How many are in the picture?
Symbol to count: t

This means you should count how many “t” there are in the image.

Task

The symbol that you will count will change in each image, so pay close attention. To make the task harder (and to prevent cheating) we have included two symbols that are very close to one another: ! and !!

These are different. So if you are asked to count ! in the image above, there are 61. If you are asked to count !!, there are only 6. Do not count !! when counting !

PLEASE NOTE:You must type the exact correct answer in order to advance to the next image. Counting each image should take about 30 seconds.

Task

This is the end of the instructions. Reminder: you will be asked questions about your willingness to complete more of this task for additional pay at the end of this initial block of work. When you click to advance to the next slide, you will begin the 5 mandatory tasks.

[Participants then completed five tasks as above]

As of right now, you have earned $1 for completing the tasks and for your overall participation in this study.

In a few minutes, we will ask you a series of questions about your willingness to do additional tasks to increase your payout.

You have already sampled the task and we will ask you about your willingness to complete more of the same tasks. The task will not change from your sample experience, except that you will have different tables to count.

We will ask you about your willingness to do additional tasks beyond those that you have already done for some additional payment.

For instance, we might ask you how many tasks you’d be willing to do for an additional $1. The exact questions we will ask you were determined in advance of your personal participation in this experiment.

In order for us to best understand your willingness to do additional tasks for additional pay, we will ask you to rank the options you see. Each option consists of some amount of tasks that you must complete and some fixed payment (in dollars). You must rank them by using the mouse and dragging the alternatives into your preferred order (where 1 is your most preferred alternative and 4 is your least preferred alternative).

In order to make sure you take each ranking seriously, we will use a particular system.

After you have ranked the alternatives, we will choose two at random. We will then implement your ranking: you will receive whichever one you assigned a higher ranking.

This may seem complicated but there is a simple and easy way to ensure you get your most-preferred outcome (subject to chance): simply answer truthfully.

Example:

Suppose you were asked to rank the following choices (this example is for illustrative purposes only):

☐ Apple
☐ Banana
☐ Orange

If you ordered the alternatives

1 Banana
2 Orange
3 Apple,

and we randomly chose (Apple, Orange), then you would get an Orange.

Quiz:

Suppose you were asked to rank Apple, Banana, Orange, and Pineapple. If your preferred ranking was Banana, Apple, Pineapple, Orange and we randomly chose (Pineapple, Apple), which would you get?

☐ Banana
☐ Orange
☐ Apple
☐ Pineapple

Correct! We will take whatever ranking you give us and implement it on the randomly selected pair of items.

Final Instructions:

Recall that we will ask you to rank different amounts of additional work and different amounts of additional pay. Your pay is on top of the money you have already earned thus far. You must complete the additional work in order to complete the experiment.

The next screen shows the real alternatives and asks for your ranking. This will count for real and will determine your payment and how much work you do, so choose carefully.

CHOICE SCREEN

[Wide money treatment; participants saw one of the two treatments listed below]

Please rank the four alternatives below:

☐ 18 tasks, receive $0.50
☐ 14 tasks, receive $2.20
☐ 18 tasks, receive $2.80
☐ 14 tasks, receive $4.50

[Wide effort treatment]

Please rank the four alternatives below:

☐ 2 tasks, receive $2.80

☐ 14 tasks, receive \$2.20
☐ 18 tasks, receive \$2.80
☐ 30 tasks, receive \$2.20

We will now randomly select two of the four alternatives that you ranked, and you will receive whichever you ranked higher.

We have randomly selected [Experimental logic chooses two at random]

and you ranked the following choice higher: [Insert choice here]

You will now complete [X] additional tasks to earn your additional pay.

Please click the continue button to complete your additional tasks. Once you have completed the additional tasks, we will ask a short series of demographic questions and the survey will end.

You will know that you have reached the end of the survey when you see a screen saying "THIS IS THE END OF THE SURVEY". Please do not exit until you have seen this screen. This screen includes the code that you must input into MTurk in order to receive payment.

[Participant completes additional tasks]

Thank you for participating. Your responses have been stored. A code to input into Amazon's MTurk is on the screen that follows. Payments will be processed within one week.

Please click the final button below to submit your work.

F Mapping the Theory to the Real-Effort Experiment

We first present a simple lemma that is useful for mapping the theory to the real-effort experiment.

Lemma F.1. *If lottery F over $\mathbb{R}$ first-order stochastically dominates lottery G over $\mathbb{R}$, then $E_F[x] + 1/2 \cdot S_F[x] > E_G[x] + 1/2 \cdot S_G[x]$ and $E_F[x] - 1/2 \cdot S_F[x] > E_G[x] - 1/2 \cdot S_G[x]$.*

Recall that participants faced the following menus of options:

Wide-Effort Treatment		*Wide-Money Treatment*	
A	(\$2.80, 2 tasks)	A'	(\$4.50, 14 tasks)
B	(\$2.20, 14 tasks)	B	(\$2.20, 14 tasks)
C	(\$2.80, 18 tasks)	C	(\$2.80, 18 tasks)
D	(\$2.20, 30 tasks)	D'	(\$0.50, 18 tasks)

We assume participants' un-normed utility for trading off effort and pay is given by $U = v(m) - \psi(e)$, where m equals dollars paid, e equals the number of tasks, and $v(\cdot), \psi(\cdot)$ are monotonically increasing functions.

We present two primary interpretations for how participants might have thought about their task. Given our elicitation method—under which participants ranked the menu of options—participants may have:

1. Viewed the choice as deterministic and reported their ranking according to $U^N(c|C)$: e.g., in the "wide-effort" treatment they ranked $A \succ B \succ C \succ D$ if $U^N(A|C^{\text{wide effort}}) > U^N(B|C^{\text{wide effort}}) > U^N(C|C^{\text{wide effort}}) > U^N(D|C^{\text{wide effort}})$.

2. Recognized that a choice of ranking was really a choice over lotteries. Given the way we incentivized rankings, ranking $A \succ B \succ C \succ D$ was, in effect, selecting a lottery where participants got paid and exerted effort according to A with probability $1/2$, according to B with probability $1/3$, and according to C with probability $1/6$. Note that this is incentive compatible with ranking options according to $U^N(\cdot|\mathscr{F})$, where $\mathscr{F}$ is the choice set of lotteries (/rankings) since options ranked higher are selected with strictly greater probability.[13]

We show below that for both of these interpretations, our theory predicts that: (i) participants will rank A [A'] first and D [D'] last in the wide-effort [wide-money] treatment and (ii) participants are more likely to rank C above B in the wide effort treatment than in the wide money treatment.[14]

Establishing (i) is easy: A and A' weakly dominate all other options in their respective menus; D and D' are weakly dominated by all other options in their respective menus. Normed utility will always rank dominating options first and dominated options last since the norming weight placed on each dimension, w_i, is strictly positive and does not vary with the particular option considered within a comparison set.

To establish (ii), we just need to show that, according to either interpretation, the weight attached to the money dimension is higher in the wide-effort treatment than the wide-money treatment and the weight attached to the effort dimension is lower in the wide-effort treatment than the wide-money treatment. Equivalently, we need to show that, under either interpretation, the range along the money dimension is lower in the wide-effort treatment than the wide-money treatment and the range along the effort dimension is higher in the wide-effort treatment than the wide-money treatment. We present each interpretation in turn.

Under the first interpretation, we have that the range along the money dimension in the wide effort treatment is $v(2.80) - v(2.20)$, while in the wide money condition it is the larger $v(4.50) - v(.50)$. The range along the effort dimension in the wide effort treatment is $\psi(30) - \psi(2)$, while in the wide money treatment it is the smaller $\psi(18) - \psi(14)$.

It takes a bit more work to establish (ii) under the second interpretation. Let $ABCD$ denote the lottery over $(v(m), -\psi(e))$ associated with ranking $A \succ B \succ C \succ D$, $DCBA$ denote the lottery over $(v(m), -\psi(e))$ associated with ranking $D \succ C \succ B \succ A$, etc. Note that:

1. The marginal distribution of $ACBD$ over $v(m)$ is first-order stochastically dominated by the

[13] We're implicitly abusing notation by writing $U^N(c|\mathscr{F})$ for the normed utility of a distribution that places probability 1 on c.

[14] There are other interpretations that involve participants taking into account the possibility that they could end the experiment at any time. This would not alter predictions (i) and (ii). We doubt participants were taking this possibility into account, however, as the experiment's incentives were heavily skewed towards fully completing the experiment. Participants were required to complete the number of tasks they committed to in order to receive pay. We found that failing to complete the experiment was rare—only eleven participants did not complete the final work block after completing the survey (including elicitation questions). Of those participants, five received the option they ranked first. We suspect these log-offs were for reasons unrelated to the realization of the lottery over tasks and pay.

marginal distribution of $A'CBD'$ over $v(m)$.

2. The marginal distribution of $DBCA$ over $v(m)$ first-order stochastically dominates the marginal distribution of $D'BCA'$ over $v(m)$.

3. The marginal distribution of $ABCD$ over $-\psi(e)$ first-order stochastically dominates the marginal distribution of $A'BCD'$ over $-\psi(e)$.

4. The marginal distribution of $DCBA$ over $-\psi(e)$ is first-order stochastically dominated by the marginal distribution of $D'CBA'$ over $-\psi(e)$.

Given that $ACBD$ ($DBCA$) gives the maximum (minimum) of the range along the money dimension in the wide-effort treatment and $A'CBD'$ ($D'BCA'$) gives the maximum (minimum) of the range along the money dimension in the wide-money treatment, points 1 and 2 together with Lemma F.1 imply that the range along the money dimension is wider in the wide-money treatment than in the wide-effort treatment. Similarly, points 3 and 4 together with Lemma F.1 imply that the range along the effort dimension is larger in the wide-effort treatment than the wide-money treatment. Accordingly, our theory predicts that participants are more likely to rank C above B in the wide-effort treatment than the wide-money treatment.

G Experiments: Supplemental Results

G.1 Participant Demographics, Experiment 1

Below we include summary statistics on participant demographics in the first experiment. Our treatments appear well balanced. We did not collect data on demographics in the second experiment.[15]

[15]The pilots indicated that there would be large heterogeneity in task completion times in the second experiment. As a result, the effective wage rate was low for some participants and we wanted to avoid asking further questions.

Table A1:
SUMMARY STATISTICS ON PARTICIPANT DEMOGRAPHICS, EXPERIMENT 1

	Drive for Discount		Survey After Purchase		Extra Work	
Variable	headphones	laptop	narrow $	wide $	narrow $	wide $
Age	38.68	37.36	38.20	37.83	37.43	38.57
	(12.49)	(11.71)	(12.32)	(11.92)	(11.15)	(12.94)
Income	3.02	2.98	2.99	3.00	3.13	2.88
	(1.009)	(1.124)	(1.114)	(1.177)	(1.134)	(1.172)
% Male	49.6	46.7	47.1	49.4	45.8	50.4
	(50.1)	(50.0)	(49.6)	(50.0)	(49.9)	(50.1)
Observations	248	244	255	237	236	256

Notes: Standard errors are in parentheses. Income is coded as a discrete variable which takes on values 1-5, corresponding to the following income brackets:
(1) Less than $15,000; (2) $15,000-$29,999; (3) $30,000-$59,999; (4) $60,000-$99,999;
(5) $100,000 or more
Each participant saw all three questions and thus summary statistics include redundancies. This information is presented here to highlight the balance of the panel.

G.2 Supplemental Results, Experiment 2

This section provides supplemental results on how participants ranked options. These results provide further evidence in support of our main hypothesis. Specifically, we show that participants are more likely to rank B above C in the wide-money treatment than in the wide-effort treatment. The full dataset is available from the publisher as per their open data policy.

Table A2:
ADDITIONAL INFORMATION ON PARTICIPANT RANKINGS, EXPERIMENT 2

	Wide Effort		Wide Money	
Ranking	Count	Percent	Count	Percent
$B \succ C$ and $A \succ D$	144	49%	173	63%
$C \succ B$ and $A \succ D$	131	45%	92	33%
$B \succ C$ or $D \succ A$	163	55%	184	67%
$C \succ B$ or $D \succ A$	150	51%	103	37%
# of Participants (Whole Sample)	294		276	

Notes: This table provides raw counts.

H Proofs

Proof of Proposition 1. For the first part, suppose each c is measured in utility units. The result trivially holds whenever $K = 1$ or $K = 2$, since the person will always choose to maximize consumption utility from a binary choice set for such K, so suppose $K \geq 3$. Let $\bar{\Delta} = \max_j \Delta_j(\{c, c'\})$ and m be a value of j satisfying $\Delta_m = \bar{\Delta}$. Further, let $\bar{c}_i = \max\{c_i, c_i'\}$ and $\underline{c}_i = \min\{c_i, c_i'\}$. Now construct c'' as follows:

- $c''_m = \underline{c}_m$
- $c''_k = \underline{c}_k + \bar{\Delta}$ for some $k \neq m$
- $c'' = \bar{c}_i - \bar{\Delta}$ for all $i \neq k, m$.

Note that c'' is not (strictly) dominated by c or c' since $c''_k \geq \bar{c}_k$.

Since $\Delta_j(C) = \bar{\Delta}$ for all j by construction, the agent will make a utility-maximizing choice from C. To complete the proof, we need to verify that this choice is in fact c', or $U(c'') \leq U(c')$:

$$\begin{aligned}\sum_i c''_i &= \underline{c}_m + \underline{c}_k + \bar{\Delta} + \sum_{i \neq k,m} (\bar{c}_i - \bar{\Delta}) \\ &\leq \sum_{i=1}^{K} \underline{c}_i + \bar{\Delta} \text{ (because } \bar{c}_i - \underline{c}_i \leq \bar{\Delta}) \\ &\leq \sum_i c'_i.\end{aligned}$$

For the second part, first consider the "if" direction. Suppose (1) holds, and let the comparison set equal $\{c, c', c''\}$, where c'' is defined such that

$$u_j(c''_j) = \begin{cases} u_j(c'_j) & \text{if } j \in A(c', c) \text{ or } j \in E(c', c) \\ -\bar{u} & \text{otherwise,} \end{cases}$$

where $\bar{u} > 0$ and $-\bar{u} < \min_k u_k(c'_k)$.

For $C = \{c, c', c''\}$ we have that

$$\begin{aligned}U^N(c'|C) - U^N(c|C) &= \sum_{i \in A(c',c)} w(\delta_i(c', c)) \cdot \delta_i(c', c) + \sum_{i \in D(c',c)} w(u_i(c_i) + \bar{u}) \cdot \delta_i(c', c) \\ &\geq \sum_{i \in A(c',c)} w(\delta_i(c', c)) \cdot \delta_i(c', c) + w\left(\min_{k \in D(c',c)} u_k(c_k) + \bar{u}\right) \sum_{i \in D(c',c)} \delta_i(c', c),\end{aligned}$$

which exceeds 0 for sufficiently large $\bar{u}$ by *N1* and (1). Since it is also true that $U^N(c'|C) - U^N(c''|C) = \sum_{i \in D(c',c)} w(u_i(c_i) + \bar{u}) \cdot (u_i(c'_i) + \bar{u}) > 0$, the person chooses c' from $\{c, c', c''\}$ when $\bar{u}$ is sufficiently large. Note that, by continuity, this argument also goes through if c'' is slightly perturbed so as not to be dominated.

For the "only if" direction, suppose condition (1) does not hold. Then, for any C containing $\{c, c'\}$,

$$\begin{aligned}U^N(c'|C) - U^N(c|C) &= \sum_{i \in A(c',c)} w(\Delta_i(C)) \cdot \delta_i(c', c) + \sum_{i \in D(c',c)} w(\Delta_i(C)) \cdot \delta_i(c', c) \\ &< \sum_{i \in A(c',c)} w(\delta_i(c', c)) \cdot \delta_i(c', c) + \sum_{i \in D(c',c)} w(\infty) \cdot \delta_i(c', c) \\ &\leq 0,\end{aligned}$$

where the first inequality follows from *N1*.

■

Lemma H.1. *For all non-degenerate distributions F with support on $[x,y]$, $y > x$, we have*

$$[E[F] - 1/2 \cdot S(F), E[F] + 1/2 \cdot S(F)] \subset [x,y].$$

Proof. We have

$$\begin{aligned}
E[F] + 1/2 \cdot S(F) &= E[F] + 1/2 \cdot \int\int |c - c'| dF(c) dF(c') \\
&= E[F] + 1/2 \cdot \int\int 2\max\{c, c'\} - (c + c') dF(c) dF(c') \\
&= E[F] + 1/2 \cdot [2E_F[\max\{c, c'\}] - 2E[F]] \\
&= E_F[\max\{c, c'\}] \\
&< y \text{ (for non-degenerate } F).
\end{aligned}$$

We can similarly establish that $E[F] - 1/2 \cdot S(F) > x$ for non-degenerate F. ■

Remark 1. The proof of Lemma H.1 establishes that $E[F] + 1/2 \cdot S(F) = E_F[\max\{c, c'\}]$, and we can similarly establish that $E[F] - 1/2 \cdot S(F) = E_F[\min\{c, c'\}]$. This provides an alternative expression for $\Delta_k(\mathscr{F})$:

$$\Delta_k(\mathscr{F}) = \max_{F \in \mathscr{F}} E_F[\max\{u_k(c_k), u_k(c'_k)\}] - \min_{F \in \mathscr{F}} E_F[\min\{u_k(c_k), u_k(c'_k)\}].$$

Proof of Proposition 2. It will be useful to recall Lemma 1 in Kőszegi and Rabin (2007): if F' is a mean-preserving spread of F and $F' \neq F$, then $S(F) < S(F')$.

For the first part of the proposition, let $\mathscr{F} = \{(F_1, F_2), (F_1 - G_1, F_2 + G_2)\}$ and $\mathscr{F}' = \{(F_1, F'_2), (F_1 - G_1, F'_2 + G'_2)\}$. Since (F_1, F_2) is chosen from $\mathscr{F}$, we have

$$U^N((F_1, F_2)|\mathscr{F}) - U^N((F_1 - G_1, F_2 + G_2)|\mathscr{F}) = w(\Delta_1(\mathscr{F})) \cdot E[G_1] - w(\Delta_2(\mathscr{F})) \cdot E[G_2] \geq 0, \tag{1}$$

where

$$\begin{aligned}
\Delta_1(\mathscr{F}) &= E[G_1] + \frac{1}{2}\left(S(F_1) + S(F_1 - G_1)\right) \\
\Delta_2(\mathscr{F}) &= E[G_2] + \frac{1}{2}\left(S(F_2 + G_2) + S(F_2)\right).
\end{aligned}$$

Since F'_2 is a mean-preserving spread of F_2 and G'_2 is a mean-preserving spread of G_2, we also have that $F'_2 + G'_2$ is a mean-preserving spread of $F_2 + G_2$, so Lemma 1 in Kőszegi and Rabin (2007) tells us that $\Delta_2(\mathscr{F}') \geq \Delta_2(\mathscr{F})$ with strict inequality whenever $F'_2 \neq F_2$ or $G'_2 \neq G_2$. Since it is also the case that $\Delta_1(\mathscr{F}') = \Delta_1(\mathscr{F})$, Equation (1) then implies that $U^N((F_1, F'_2)|\mathscr{F}') - U^N((F_1 - G_1, F'_2 +$

$G_2')|\mathscr{F}') \geq 0$ by *N1*, with strict inequality whenever $F_2' \neq F_2$ or $G_2' \neq G_2$.

It remains to show the second part of the proposition. Let $\mathscr{F}(G_1, G_2)$ denote the comparison set when the decision-maker faces the distribution over choice sets of the form $\{(0,0), (-\tilde{x}, \tilde{y})\}$ that is induced by drawing $\tilde{x}$ from G_1 and $\tilde{y}$ independently from G_2, where $G_1 \in \{F_1, F_1'\}$ and $G_2 \in \{F_2, F_2'\}$.

The range on each dimension equals the range when we restrict attention to the subset of $\mathscr{F}(G_1, G_2)$ generated by the union of the lotteries associated with "always choose $(0,0)$" and "always choose $(-\tilde{x}, \tilde{y})$". The first of these lotteries yields $E_F[u_k(c_k)] \pm \frac{1}{2} S_F[u_k(c_k)] = 0$ along each dimension, while the second yields $-E[G_1] \pm 1/2 \cdot S(G_1)$ along the first and $E[G_2] \pm 1/2 \cdot S(G_2)$ along the second dimension.

By Lemma H.1, the range on the dimensions are then

$$\Delta_1(\mathscr{F}(G_1, G_2)) = E[G_1] + \frac{1}{2} S(G_1)$$
$$\Delta_2(\mathscr{F}(G_1, G_2)) = E[G_2] + \frac{1}{2} S(G_2).$$

Consequently, $\Delta_2(\mathscr{F}(F_1, F_2)) < \Delta_2(\mathscr{F}(F_1, F_2'))$ whenever (i) $F_2' \neq F_2$ is a mean-preserving spread of F_2, as, in this case, $E[F_2] = E[F_2']$ and $S(F_2') > S(F_2)$ by Lemma 1 in Kőszegi and Rabin (2007), or (ii) F_2' first order stochastically dominates F_2, as, in this case, $E[F_2'] + 1/2 \cdot S(F_2') = E_{F_2'}[\max\{\tilde{y}, \tilde{y}'\}] > E_{F_2}[\max\{\tilde{y}, \tilde{y}'\}] = E[F_2] + 1/2 \cdot S(F_2)$, where the equality comes from Remark 1 and the inequality is obvious. From (i) and (ii), $\Delta_2(\mathscr{F}(F_1, F_2)) < \Delta_2(\mathscr{F}(F_1, F_2'))$ whenever $F_2' \neq F_2$ first order stochastically dominates a mean-preserving spread of F_2.

The result then follows from the fact that

$$U^N((0,0)|\mathscr{F}) - U^N((-x, y)|\mathscr{F}) = w(\Delta_1) \cdot x - w(\Delta_2) \cdot y$$

is increasing in Δ_2 by *N1*. ■

Proof of Proposition A.1. Let $d(c', c|C) \in \mathbb{R}^K$ denote a vector that encodes proportional differences with respect to the range of consumption utility: For all k,

$$d_k(c', c|C) = \frac{\delta_k(c', c)}{\Delta_k(C)}.$$

We have

$$U^N(c|C) - U^N(c'|C) = \sum_j w(\Delta_j(C)) \left[u_j(c_j) - u_j(c_j')\right] = \sum_j w\left(\frac{\delta_j(c, c')}{d_j(c, c'|C)}\right) \delta_j(c, c') \geq 0, \quad (2)$$

where the inequality follows from the person being willing to choose c from C.

For part 1, suppose $\tilde{C}$ is a k-widening of C with $\delta_k(\tilde{c},\tilde{c}') > \delta_k(c,c') > 0$, $d_k(\tilde{c},\tilde{c}'|\tilde{C}) = d_k(c,c'|C)$, and $\delta_i(\tilde{c},\tilde{c}') = \delta_i(c,c')\ \forall i \neq k$. Then

$$\begin{aligned} U^N(\tilde{c}|\tilde{C}) - U^N(\tilde{c}'|\tilde{C}) &= U^N(c|C) - U^N(c'|C) + \left[w\left(\frac{\delta_k(\tilde{c},\tilde{c}')}{d_k(c,c'|C)}\right)\delta_k(\tilde{c},\tilde{c}') - w\left(\frac{\delta_k(c,c')}{d_k(c,c'|C)}\right)\delta_k(c,c')\right] \\ &> U^N(c|C) - U^N(c'|C) \text{ (by } N2\text{)}, \end{aligned}$$

so the person is not willing to choose $\tilde{c}'$ from $\tilde{C}$.

For part 2, suppose $\tilde{C}$ is a k-widening of C with $\delta_k(c,c') < 0$ and $\delta_i(\tilde{c},\tilde{c}') = \delta_i(c,c')\ \forall i$. Then

$$\begin{aligned} U^N(\tilde{c}|\tilde{C}) - U^N(\tilde{c}'|\tilde{C}) &= U^N(c|C) - U^N(c'|C) + \left[w\left(\frac{\delta_k(c,c')}{d_k(\tilde{c},\tilde{c}'|C)}\right)\delta_k(c,c') - w\left(\frac{\delta_k(c,c')}{d_k(c,c'|C)}\right)\delta_k(c,c')\right] \\ &> U^N(c|C) - U^N(c'|C) \text{ (by } N1\text{)}, \end{aligned}$$

so the person is not willing to choose $\tilde{c}'$ from $\tilde{C}$.

$\blacksquare$

Proof of Proposition A.2. First, consider the case where c'' spreads out the advantages of c' relative to c. In this case, there exists a $j \in E(c',c)$, $k \in A(c',c)$, and $\varepsilon < \delta_k(c',c)$ such that

$$\begin{aligned} U^N(c''|\{c,c''\}) - U^N(c|\{c,c''\}) = U^N(c'|\{c,c'\}) - U^N(c|\{c,c'\}) \\ + w(\delta_k - \varepsilon)\cdot(\delta_k - \varepsilon) + w(\varepsilon)\cdot\varepsilon - w(\delta_k)\cdot\delta_k. \end{aligned}$$

Supposing that $\varepsilon \leq \frac{\delta_k(c',c)}{2}$ (the case where $\frac{\delta_k(c',c)}{2} < \varepsilon < \delta_k(c',c)$ is analogous), the result then follows from the fact that

$$w(\delta_k - \varepsilon)\cdot(\delta_k - \varepsilon) + w(\varepsilon)\cdot\varepsilon - w(\delta_k)\cdot\delta_k \geq w(\delta_k - \varepsilon)\cdot\delta_k - w(\delta_k)\cdot\delta_k > 0,$$

by successive applications of *N1*.

Now consider the case where c'' integrates the disadvantages of c' relative to c. In this case, there exists $j,k \in D(c',c)$ such that $U^N(c''|\{c,c''\}) - U^N(c|\{c,c''\})$ equals

$$\begin{aligned} U^N(c'|\{c,c'\}) - U^N(c|\{c,c'\}) + w\left(|\delta_j(c',c) + \delta_k(c',c)|\right)\cdot(\delta_j(c',c) + \delta_k(c',c)) \\ - \left[w(|\delta_j(c',c)|)\cdot\delta_j(c',c) + w(|\delta_k(c',c)|)\cdot\delta_k(c',c)\right]. \end{aligned}$$

The result then follows from the fact that

$$w\left(|\delta_j(c',c) + \delta_k(c',c)|\right)\cdot(\delta_j(c',c) + \delta_k(c',c)) > w(|\delta_j(c',c)|)\cdot\delta_j(c',c) + w(|\delta_k(c',c)|)\cdot\delta_k(c',c),$$

by *N1* (recall that $\delta_i(c',c) < 0$ for $i = j,k$). ■

Proof of Proposition A.3. From Condition (1) of Proposition 1 we want to show that

$$\sum_{i \in A(c'',c)} w(\delta_i(c'',c)) \cdot \delta_i(c'',c) + \sum_{i \in D(c'',c)} w(\infty) \cdot \delta_i(c'',c) > 0.$$

Since there is a C containing $\{c,c'\}$ such that c' is chosen from C, Condition (1) must hold for c',c. Noting that $\sum_{i \in D(c'',c)} w(\infty) \cdot \delta_i(c'',c) = \sum_{i \in D(c',c)} w(\infty) \cdot \delta_i(c',c)$, it suffices to show that

$$\sum_{i \in A(c'',c)} w(\delta_i(c'',c)) \cdot \delta_i(c'',c) \geq \sum_{i \in A(c',c)} w(\delta_i(c',c)) \cdot \delta_i(c',c),$$

which can be shown via an argument analagous to the one we made in the proof of Proposition A.2. ■

Proof of Corollary A.1. Let $d(c',c|C) \in \mathbb{R}^K$ denote a vector that encodes proportional differences with respect to the range of consumption utility: For all k,

$$d_k(c',c|C) = \frac{\delta_k(c',c)}{\Delta_k(C)}.$$

1. If c dominates c', then $D(c',c)$ is non-empty, while $A(c',c)$ is empty, implying that condition (1) does not hold. The result then follows from Proposition 1.

2. Fix $\tilde{\delta}_A$. The left-hand side of condition (1) equals

$$\sum_{i \in A(c',c)} w(\delta_i(c',c)) \cdot \delta_i(c',c) - w(\infty) \cdot \delta_D(c',c) \leq \left\{ \sup_{\{\mathbf{d} \in \mathbb{R}_+^K : \sum_i d_i = \tilde{\delta}_A\}} \sum_{i=1}^{K} w(d_i) \cdot d_i \right\} - w(\infty) \cdot \delta_D(c',c).$$

   Clearly, the right-hand side of the above inequality falls below 0 for $\delta_D(c',c)$ sufficiently large when *N3* holds. The result then follows from Proposition 1.

■

Proof of Proposition A.4. The result is trivial for $K = 1$, so suppose $K \geq 2$. Let $C_\varepsilon = \{c\} \cup \{c^1\} \cup \ldots \cup \{c^K\}$, where, for each $j \in \{1,\ldots,K\}$, define $c^j \in \mathbb{R}^K$ such that

$$u_i(c_i^j) = \begin{cases} u_i(c_i) & \text{for } i = j \\ u_i(c_i) - \bar{u} & \text{for all } i \neq j, \end{cases}$$

supposing $\bar{u}$ is sufficiently large that $w(\bar{u}) - w(\infty) < w(\infty) \cdot \varepsilon \equiv e$.

By construction, $\Delta_i(C_\varepsilon) = \bar{u}$ for all i, so the person makes a utility maximizing choice from C_ε, which is c.

Also, for any $\tilde{C}$ containing C_ε, $\Delta_i(\tilde{C}) \geq \Delta_i(C_\varepsilon)$, so $w(\Delta_i(\tilde{C})) < w(\infty) + e$ for all i. This means that, for any $c' \neq c \in \tilde{C}$, we have

$$
\begin{aligned}
U^N(c'|\tilde{C}) - U^N(c|\tilde{C}) &= \sum_{i \in A(c',c)} w(\Delta_i(\tilde{C})) \cdot \delta_i(c',c) + \sum_{i \in D(c',c)} w(\Delta_i(\tilde{C})) \cdot \delta_i(c',c) \\
&< (w(\infty) + e) \cdot \delta_A(c',c) - w(\infty) \cdot \delta_D(c',c),
\end{aligned}
$$

where this last term is negative (meaning c' will not be chosen from $\tilde{C}$) whenever $\delta_A(c',c) = 0$ or $\delta_A(c',c) > 0$ and $\frac{\delta_D(c',c)}{\delta_A(c',c)} - 1 > \varepsilon$. ■

Proof of Corollary A.2. This result is trivial if $K = 1$ or $\delta_A(c',c) = 0$, so let $K \geq 2$ and $\delta_A(c',c) > 0$. Since $U(c) > U(c')$,

$$
\lambda \equiv \frac{\delta_D(c',c)}{\delta_A(c',c)} - 1 > 0.
$$

Let $C = \{c'\} \cup C_{\lambda'}$, where $C_{\lambda'}$ is constructed as in the proof of Proposition A.4 (letting $\varepsilon = \lambda'$), with $\lambda' = \lambda - \eta$ for $\eta > 0$ small.

By Proposition A.4, c would be chosen from $C_{\lambda'}$ and c' would not be chosen from any $\tilde{C}$ containing $C_{\lambda'}$ and c', including from C. It is left to establish that c would be chosen from C, but this follows from the fact that $U^N(c|C) - U^N(c^j|C) = \sum_{i \neq j} w(\Delta_i(C)) \cdot \bar{u} \geq 0$ for any $c^j \in C_{\lambda'}$. ■

Proof of Lemma F.1. We know from Lemma H.1 that $E_F[x] + 1/2 \cdot S_F[x] = E_F[\max\{x,x'\}]$, which is larger than $E_G[\max\{x,x'\}]$ since F first-order stochastically dominates G. (If F is the distribution over x then $F(x)^2$ is the distribution over $\max\{x,x'\}$ so the first-order stochastic dominance relation is preserved when taking first-order statistics.) Likewise, from Lemma H.1, $E_F[x] - 1/2 \cdot S_F[x] = E_F[\min\{x,x'\}]$, which is greater than $E_G[\min\{x,x'\}]$ since F first-order stochastically dominates G. (If F is the distribution over x then $1 - (1 - F(x))^2$ is the distribution over $\min\{x,x'\}$, which is increasing in $F(x)$, so the first-order stochastic dominance relation is preserved when taking second-order statistics.) ■

References

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice Under Risk," *The Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.

_ , _ , and _ , “Salience and Consumer Choice,” *Journal of Political Economy*, 2013, *121* (5), 803–843.

Cunningham, Tom, “Comparisons and Choice,” 2013. Mimeo, Stockholm University.

Koszegi, Botond and Adam Szeidl, “A model of focusing in economic choice,” *The Quarterly Journal of Economics*, 2013, *128* (1), 53–104.

Parducci, Allen, “Category Judgment: A Range-Frequency Model.,” *Psychological Review*, 1965, *72* (6), 407 – 418.

Thaler, Richard, “Mental Accounting and Consumer Choice,” *Marketing Science*, 1985, *4* (3), 199–214.

Thaler, Richard H, “Mental Accounting Matters,” *Journal of Behavioral Decision Making*, 1999, *12*, 183–206.

_ and Eric J Johnson, “Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice,” *Management science*, 1990, *36* (6), 643–660.