A Model of Relative Thinking

Benjamin Bushong
Harvard

Matthew Rabin
Harvard

Joshua Schwartzstein*
Harvard Business School

March 31, 2016

Abstract

Fixed differences loom smaller when compared to large differences. We propose a model of relative thinking where a person weighs a given change along a consumption dimension by less when it is compared to bigger changes along that dimension. In deterministic settings, the model predicts context effects such as the attraction effect, but predicts meaningful bounds on such effects driven by the intrinsic utility for the choices. In risky environments, a person is less likely to exert effort in a money-earning activity if he had expected to earn higher returns or if there is greater income uncertainty. In intertemporal consumption, relative thinking induces a tendency to overspend and for a person to act more impatient if infrequently allotted large amounts to consume than if frequently allotted a small amount to consume, or especially the greater the uncertainty in future consumption utility.

Keywords: relative thinking; JEL Classification: B49

*E-mails: bushong@fas.harvard.edu, matthewrabin@fas.harvard.edu, and jschwartzstein@hbs.edu. We are grateful to Roland Bénabou, Pedro Bordalo, Tom Cunningham, Erik Eyster, Drew Fudenberg, Xavier Gabaix, Tristan Gagnon-Bartsch, Botond Kőszegi, Kristóf Madarász, Jesse Shapiro, Andrei Shleifer, Dmitry Taubinsky, and seminar participants at BEAM, Berkeley, ESSET, Harvard, LSE, Princeton, and Tennessee for helpful feedback. We also thank Xiaosheng Mu for excellent research assistance.

"All brontosauruses are thin at one end, much, *much* thicker in the middle, and then thin again at the far end."

— Ms. Anne Elk's theory of the brontosaurus, from a *Monty Python* sketch

1 Introduction

The quote above is from a lesser-known *Monty Python* sketch, portraying an annoying scientist taking a very long time to reveal her theory of the brontosaurus, which ultimately consisted solely of the above quote. Although we may be underwhelmed by the empirical insight, we all believe her theory is obviously right. Yet in what sense is this so? The tapered shape of the brontosaurus we all picture is thin at the ends *relative* to the middle. But not in *absolute* terms: the brontosaurus's front and back are surely thicker in absolute terms than a frog's—even though we would not describe the frog as being thin at the ends. "Thin" is a relative term. Amounts appear smaller when compared to larger things than when compared to smaller things.

But quite what we mean by "relative thinking" is not entirely clear in either perception or—more importantly for our purposes—in economic choice. Indeed, alternative models channel relative thinking in different ways. Cunningham (2013) outlines an approach building from the idea that large average magnitudes along a dimension, measured relative to zero, make given-sized items loom less large than they would if the average on that dimension were smaller. The oft-invoked notion of "diminishing sensitivity", built into Kahneman and Tversky (1979) and more recently into choice-set dependent models like Bordalo, Gennaioli, and Shleifer (2012, 2013), says that an incremental change matters less the further it is from a zero point. This paper captures an additional way relative thinking matters for choice that has long been a pervasive psychological intuition. Starting with at least Volkmann (1951) and Parducci (1965), research has explored how a given absolute difference can seem big or small depending on the *range* under consideration. Following papers such as Mellers and Cooke (1994), and more recently Soltani, De Martino, and Camerer (2012), we study how the range of alternatives under consideration affects choice. But we broaden the domain of application to choice under uncertainty and intertemporal choice, and flesh out the assumptions necessary to embed range-based relative thinking into economic models.

The model of range-based relative thinking this paper develops assumes a person puts less weight on incremental changes along a consumption dimension when outcomes along that dimension exhibit greater variability in the choice set he faces. We draw out the model's general features and implications and compare our "range-based" approach to other related concepts. The model does well at matching known context effects from psychology and marketing, for example proportional thinking (Thaler 1980, Tversky and Kahneman 1981) and attraction effects (Huber, Payne, and

Puto 1982). But in contrast to the selective-illustration approach of much of the literature on context effects, we provide a more complete characterization of the effects of range-based relative thinking on choice across contexts. In the process, we clarify limits to context effects implied by the model, showing for instance necessary and sufficient conditions for there to exist *no* context that induces a person to choose a particular lower-utility option over a particular higher-utility option. Our primary emphasis, however, is on the model's economic implications. In the context of discretionary labor choices, for example, the model says that a worker will choose to exert less effort for a fixed return when there is greater overall income uncertainty. We also develop predictions of a variant of our model about spending patterns over time, showing why relative thinking induces a tendency to overspend. A relative thinker on a limited budget appears more impatient the longer the horizon of consumption, (roughly) when he is less frequently allotted larger amounts to consume, and especially the greater the uncertainty in future consumption utility.

Given the multitude of intuitions about how relative thinking influences choice, consider a few examples which highlight how the act of comparison naturally invokes a form of range-based relative thinking. If a traveling parent assesses the size of toy dinosaurs at the airport gift shop to bring home to her kids, or lattes at the airport Starbucks to keep her awake, she compares sizes among options. Adding a much bigger toy dinosaur to a pair of slightly different dinosaurs is likely to make them both look smaller, and also make the difference between them look smaller. But instead adding a much *smaller* dinosaur to this same pair might have us perceive them as larger—but *once again* make the difference between them look smaller. Adding *both* smaller and larger dinosaurs to the set may not induce a clear shift in their perceived size (even Ann Elk might get confused), but surely will make the difference between them look smaller. More generally, there are all sorts of situations where two different choice sets differ in range while fixing various notions of the zero points used to judge distance when channeling diminishing sensitivity (variously the low point, the average, the expected choice, or the historical choice). Likewise there are many cases where the size of the range of choices can be disentangled from notions of the magnitude under consideration (variously the highest point or the average). The intuitive and much researched perspective that wider ranges make people less attentive to the incremental differences has been missing from economic analysis.

And we believe such range-based relative thinking matters. It will matter at the airport for small decisions (toys for a child) and big decisions (caffeine for yourself). But it can also matter for very big decisions. If you are shopping for houses, seeing a big house may make a pair of small houses you have seen look smaller; likewise, seeing a small house may make a pair of large houses seem larger. But we think adding that third house will, by making the range of houses so much bigger, decrease the perceived difference in the original pair and reduce the person's willingness to pay for slight improvements in size among that pair. Likewise, if a person is considering a set of different possible jobs that range narrowly in salary between \$108,000 and \$113,000, we believe the \$5,000 salary differential would loom larger in the choice than that \$5,000 differential for the same two jobs would loom if the ranges were instead set by \$74,000-\$114,000, \$104,000-\$192,000, or \$68,000-\$153,000. Or consider what happens with uncertainty. If you are searching for houses and uncertain a priori what the prices or sizes of houses will be, our model says that the wider the variance (using a particular measure), the wider the "range" and the less you will attend to small differences in prices or sizes. And, related to an example we return to, our model says that the more uncertain an Uber driver is about daily earnings—realizing it could be very high or very low—the less motivated he will be to put in effort to earn an extra \$10.

We present our model for deterministic environments in Section 2. A person's "consumption utility" for a K-dimensional consumption bundle c is separable across dimensions: $U(c) = \sum_k u_k(c_k)$. We assume that a person instead makes choices according to "normed consumption utility" that depends not only on the consumption bundle c but also the comparison set C—which in applications we will equate with the choice set. To do so, we follow a recent economic literature begun by Bordalo, Gennaioli, and Shleifer (2012) in assuming that the comparison set influences choice through distorting the relative weights a person puts on consumption dimensions. Normed consumption utility equals $U^N(c|C) = \sum_k w_k \cdot u_k(c_k)$, where w_k captures the weight that the person places on consumption dimension k given the (notationally suppressed) comparison set C. The weights $w_k > 0$ are assumed to be a function $w_k \equiv w(\Delta_k(C))$, where $\Delta_k(C) = \max_{\tilde{c} \in C} u_k(\tilde{c}_k) - \min_{\tilde{c} \in C} u_k(\tilde{c}_k)$ denotes the range of consumption utility along dimension k. Our key assumption is that $w(\Delta)$ is decreasing: the wider the range of consumption utility on some dimension, the less a person cares about a fixed utility difference on that dimension.[1] For example, in searching for flights on a flight aggregator like Orbitz, the model says that spending extra for convenience will feel bigger when the range of prices is \$250 - \$450 than when the range is \$200 - \$800. We also assume that $w(\Delta) \cdot \Delta$ is increasing, so that differences in normed utility are increasing in absolute magnitude when fixed as a proportion of the range: The \$600 difference

[1] A similar theme has been heavily emphasized in recent neuroscience. Models of normalization, such as the notion of "range-adaptation" in Padoa-Schioppa (2009) or "divisive normalization" in Louie, Grattan and Glimcher (2011), tend to relate both the logic of neural activity, and the empirical evidence (reviewed in Rangel and Clithero 2012) on the norming of "value signals", to the possible role of norming in simple choices. Fehr and Rangel (2011) argue that "the best and worst items receive the same decision value, regardless of their absolute attractiveness, and the decision value of intermediate items is given by their relative location in the scale." Insofar as these models of value coding in the brain translate into values driving economically interesting choice, they may provide backing to the ideas discussed in this paper. Mellers and Cooke (1994) provide experimental evidence that trade-offs more generally depend on attribute ranges, where the impact of a fixed attribute difference is larger when presented in a narrower range. There is also a related marketing literature (see, e.g., Janiszewski and Lichtenstein 1999) on "price perceptions", which presents suggestive evidence that whether a given price seems big or small depends on its relative position in a range. Soltani, De Martino and Camerer (2012), which we discuss below, provide behavioral evidence consistent with range-based normalization being responsible for classical decoy effects, like the "attraction effect". Indeed, range-frequency theory (Parducci 1974) motivated Huber, Payne and Puto's (1982) original demonstrations of such effects.

seems bigger when the range is \$600 than the \$200 difference seems when the range is \$200. Finally, we assume that $w(\Delta)$ is bounded away from zero: large differences cannot loom arbitrarily small no matter how big the ranges they are compared to.

Section 3 explores context effects in riskless choice that are induced by relative thinking. When a person starts off indifferent between two 2-dimensional alternatives, the addition of a third to the comparison set influences his choices in ways consistent with experimental evidence. For example, our model predicts the "asymmetric dominance effect" proposed by Huber, Payne and Puto (1982): Adding a more extreme, inferior option to the choice set leads the person to prefer the closer of the two superior options, since the addition expands the range of the closer option's disadvantage relative to the other alternative by more than it expands the range of its advantage. Section 3 also characterizes some implications of our model for the limits of context effects based on bounds placed on the weighting function. Some inferior options, even if undominated, can never be "normed" into selection.

Section 4 extends the model to choice under uncertainty, letting us study range effects in both choices over lotteries and in situations where a person makes plans prior to knowing the exact choice set he will face. The key assumption in extending our model to risky choice is that it is not only the range of expected values across lotteries that matters, but also the range of outcomes in the support of given lotteries. Roughly, we summarize each lottery's marginal distribution over $u_k(c_k)$ in terms of its mean plus or minus a measure of its variation, and take the range along a dimension to equal the difference between the maximal mean-plus-variation among all lotteries, and the minimal mean-minus-variation.

In Section 5, we spell out the implications of our model in two economic contexts. First, we flesh out direct implications of the Section 4 uncertainty model: People are more inclined to sacrifice on a dimension when it is riskier. For example, people are less willing to put effort into a money-earning activity when either a) they earn money simultaneously from another stochastic source, or b) they faced a wider range of *ex ante* possible returns; in both cases, the wider range in the monetary dimension due to the uncertainty lowers their sensitivity to incremental changes in money. Likewise, in these cases the wider range makes effort choices less sensitive to the level of monetary incentives. Finally, people are less willing to put in effort for a fixed return when they expected the opportunity to earn more because this also expands the range on the monetary dimension and makes the fixed return feel small.

Section 5 also extends the model to consider how a person trades off consumption across time. Applying the model to this environment requires assumptions on the degree to which the person thinks about consumption at different points of time as consumption across different dimensions, as well as on how he values any money left over for tomorrow from today's perspective. We suppose that the person segregates out consumption today, but integrates consumption in future periods together into one dimension and values money left over for tomorrow as if he optimally spends it to maximize consumption utility. Intuitively, the person thinks more precisely about how he spends money today than tomorrow. Under these assumptions, the range of future consumption utility tends to be larger than the range of current consumption utility, so that relative thinking induces the person to act present-biased. More novelly, the degree of expressed present bias is increasing in factors that magnify how much bigger the range of future consumption utility is than the range of current consumption utility, including the longer the horizon of consumption and the greater the uncertainty in future consumption utility. The model also implies that when a person has a fixed amount of money to spend on consumption over a finite number of periods then he is more likely to spend on a good if he can buy it at the beginning, even if he has to wait to consume the good.

The notion of proportional thinking that is inherent in range-based relative thinking is a frequent motivator for the idea that people exhibit diminishing sensitivity to changes the further those changes are from a reference point. Range-based relative thinking is different: In the presence of greater ranges along a dimension, our model says *all* changes along that dimension loom smaller. When considering possibilities of large-scale decisions, smaller stakes seem like peanuts. Many classical demonstrations of diminishing sensitivity are confounded with the effects of range-based relative thinking: when zero is natural as both an end point and the reference point, both relative thinking and diminishing sensitivity say that the marginal dollar feels larger at \$13 when \$13 is the maximal loss than the marginal dollar does at \$283 when *it* is the maximal loss. But diminishing sensitivity says that the difference between losing \$12 vs. \$13 feels bigger than the difference between losing \$282 vs. \$283, irrespective of the range of potential losses. Range-based relative thinking predicts the two will feel the same for any fixed range—but, importantly for generating context effects such as the attraction effect and for predicting reduced sensitivity in the presence of uncertainty, the \$12 vs. \$13 difference will loom larger when \$13 is the maximal potential loss than when \$283 is the maximal potential loss. While diminishing sensitivity is independently a real phenomenon in both the psychophysics of perception and in choice behavior, range effects are a distinct and important influence on perception and choice, and we feel that some instances of relative thinking have been mistaken for diminishing sensitivity.

Because we posit that features of the choice context influence how attributes of different options are weighed, at a basic level our model relates to Bordalo, Gennaioli, and Shleifer's (2012, 2013) approach to studying the role of salience in decision-making, as well as more recent models by Kőszegi and Szeidl (2013) and Cunningham (2013). We briefly compare the predictions of these models to ours in Section 3, and Appendix C shows in more detail how each of these models generates substantively different predictions to ours in some environments, because none universally share our property that fixed differences along a dimension loom smaller in the presence of bigger ranges.[2] In many ways, the framework for our riskless model most closely resembles Kőszegi and Szeidl's (2013) model of focusing, and indeed elements of our formalism build directly from it.[3] But it sits in an interesting and uncomfortable relationship to their model: we say the range in a dimension has the exact opposite effect as it does in their model. Although some of their examples—as well as those in Bordalo, Gennaioli, and Shleifer (2013)—are compelling about how attentional and focusing issues can lead wider ranges to enhance the weight a person places on a dimension, we cannot share their intuition that such effects dominate relative-thinking effects in most of the two-dimensional economic-choice situations that we focus on in this paper. We sketch a framework for studying the interaction between focusing effects and relative thinking in Section 6. We conclude in Section 7 by discussing normative issues and shortcomings of the model.

2 Relative Thinking: The Deterministic Case

We begin by presenting a special case of the model that applies to a situation where a person chooses from sets of riskless options. In later sections we present the full model which also enables us to study choice under uncertainty, defining ranges as a function of available lotteries. The agent's "hedonic utility" for a riskless outcome is $U(c) = \sum_k u_k(c_k)$, where $c = (c_1, \ldots, c_K) \in \mathbb{R}^K$ is consumption and we assume each $u_k(c_k)$ is strictly increasing in c_k. The person does not maximize hedonic utility, but rather "normed" utility, which for a riskless option equals $U^N(c|C)$ given "comparison set" C.[4]

[2] Azar (2007) provides a theory of relative thinking built on diminishing sensitivity, where people are less sensitive to price changes at higher price levels. Contemporaneously, Kontek and Lewandowski (2013) present a model of range-dependent utility to study risk preferences in choices over single-dimensional lotteries. They assume that the outcomes of a given lottery are normed only according to the range of outcomes within the support of that lottery—other lotteries in the choice set do not influence a lottery's normed utility. Thus, unlike our model, theirs is not one of context-dependent choice or valuation. Less closely related, Tversky (1969) considers a model of intransitive preferences where, in binary decisions, people neglect a dimension when the difference between alternatives on that dimension is sufficiently small. Rubinstein (1988) proposes a related model of how people choose between lotteries which vary in the probability and magnitude of lottery prizes that can account for the Allais paradoxes.

[3] Kőszegi and Szeidl's (2013) model does not study uncertainty; extending their model along the lines of our extension in Section 4 would be straightforward. The underlying psychology of our model is more closely related to Cunningham (2013). However, since Cunningham (2013) models proportional thinking in relation to the average size of attributes rather than as a percentage of the range, predictions depend on the choice of a reference point against which the size of options is defined. In our understanding of Cunningham (2013) as positing zero as the reference point, a person would be *less* sensitive to paying \$1200 rather than \$1100 for convenience if the choices ranged between \$1100 and \$1200 than if they ranged between \$400 and \$1300. Our model says the narrower range would make people more sensitive; models of diminishing sensitivity would say it would not matter (fixing the reference point). In this sense, while the motivation of Cunningham (2013) is the most similar to our model and a precursor to our model, we feel the average-level-based formalization captures something very different than our ranged-based formalization of relative thinking.

[4] Although we do not emphasize normative implications through most of the paper, the labeling here of the unnormed utility as "hedonic" connotes our perspective that norming may influence choice without affecting experienced

Throughout this paper, we equate the comparison set with the (possibly stochastic) choice set, though the model setup is more general.[5] When a person makes plans prior to knowing which choice set he will face, options outside the *realized* choice set can matter. We will more carefully describe how this works when we consider choices over lotteries in Section 4. Our model assumes that the comparison set influences choice through distorting the relative weight a person puts on each consumption dimension. Normed consumption utility equals

$$U^N(c|C) = \sum_k u_k^N(c_k|C) = \sum_k w_k \cdot u_k(c_k),$$

where w_k captures the weight that the decision-maker places on consumption dimension k given comparison set C.[6]

We make the following assumptions on w_k:

Norming Assumptions in the Deterministic Case:

N0(d). The weights w_k are given by $w_k = w(\Delta_k(C))$, where $\Delta_k(C) = \max_{c \in C} u_k(c_k) - \min_{c \in C} u_k(c_k)$ denotes the range of consumption utility along dimension k.

N1. $w(\Delta)$ is a differentiable, decreasing function on $(0, \infty)$.

N2. $w(\Delta) \cdot \Delta$ is defined on $[0, \infty)$ and is strictly increasing.

The first two assumptions capture the psychology of relative thinking: the decision-maker attaches less weight to a given change along a dimension when the range of consumption utility along that dimension is higher. Put differently, these assumptions imply that a particular advantage or disadvantage of one option relative to another looms larger when it represents a greater percentage of the overall range.[7] Although one could use notions of dispersion that rely on more

utility.

[5]Equating choice and comparison sets gives us fewer degrees of freedom and allows us to make predictions based solely on the specified probability distribution over choice sets. In many situations, we think it is realistic to exclude options from the comparison set when they lie outside the choice set. There are exceptions, however. For example, as discussed by Bordalo, Gennaioli, and Shleifer (2015) and Cunningham (2013), options faced in the past might enter into the comparison set even when people cannot plausibly attach positive probability to facing these options again in the future.

[6]While this formulation is sufficient for analyzing choice behavior in the way we do, as we discuss in Section 7 this formulation could be inadequate to do cross-choice-set welfare analysis. In that context, we could instead consider mathematically re-normalized formulations such as

$$\tilde{U}^N(c|C) = \sum_k \min_{\tilde{c} \in C} u_k(\tilde{c}_k) + w_k \cdot (u_k(c_k) - \min_{\tilde{c} \in C} u_k(\tilde{c}_k)).$$

This re-normalization to $\min_{\tilde{c} \in C} u_k(\tilde{c}_k)$ may provide a more natural interpretation across contexts, since it implies that normed and un-normed decision utility coincide given singleton comparison sets.

[7]Even if one somehow found natural units to choose within a dimension, the natural unit of comparison is util-

than just the endpoints, we follow Parducci (1965), Mellers and Cooke (1994), and Kőszegi and Szeidl (2013) by assuming that the range of consumption utility in the deterministic case (the "d" in *N0(d)* stands for deterministic) is simply the difference between the maximum value and the minimum value. In later sections we present the full assumption *N0* that handles lotteries.

Assumption *N2* assures that people are sensitive to absolute consumption utility differences. If a person likes apples more than oranges, then he strictly prefers choosing an apple when the comparison set equals $\{(1 \text{ apple}, 0 \text{ oranges}), (0 \text{ apples}, 1 \text{ orange})\}$: While *N1* says that the decision weight on the "apple dimension" is lower than the decision weight on the "orange dimension"—since the range of consumption utility on the apple dimension is higher—*N2* guarantees that the trade-off between the two dimensions still strictly favors picking the apple. In particular, giving up 100% the range on the apple dimension looms strictly larger than gaining 100% the range on the orange dimension. This assumption is equivalent to assuming that the decision weight is not too elastic with respect to the range. In the limiting case where $w(\Delta) \cdot \Delta$ is constant in Δ, the agent only considers percentage differences when making decisions.

We sometimes add a final assumption, which bounds the impact of relative thinking:

N3. $\lim_{\Delta \to \infty} w(\Delta) \equiv w(\infty) > 0$ and $w(0) < \infty$.

This assumption says that a given difference in consumption utility is never negated by norming, and arbitrarily large differences are arbitrarily large even when normed. Similarly, it says a given difference in consumption utility is never infinitely inflated by norming, and arbitrarily small differences are arbitrarily small even when normed. While we assume that *N0-N2* hold throughout the paper, we specifically highlight when results rely on *N3*, since the limiting behavior of $w(\cdot)$ only matters for a subset of the results and we view this assumption as more tentative than the others.

The notation and presentation implicitly build in an important assumption: The weight on a dimension depends solely on the utility range in that dimension. The universal $w(\Delta)$ function implies a no-degree-of-freedom prediction once a cardinal specification of utilities is chosen—with the important restriction to dimension-separable utility functions.[8]

In Appendix B we discuss a method for determining both $u_k(\cdot)$ and $w(\cdot)$ from behavior, which closely follows the approach in Kőszegi and Szeidl (2013). The elicitation assumes that we know

ity rather than consumption levels given our interest in tradeoffs across dimensions. In terms of capturing the psychophysics, using utility may miss neglect of diminishing marginal utility: a person faced with 100 scoops of ice cream may treat the difference between 2-3 scoops as smaller than if he faced the possibility of getting 5 scoops, even though he may be satiated at 5 scoops.

[8] Once $w(\cdot)$ is fixed, affine transformations of $U(\cdot)$ will not in general result in affine transformations of the normed utility function. As such, like other models that transform the underlying "hedonic" utility function, either $U(\cdot)$ must be given a cardinal interpretation or the specification of $w(\cdot)$ must be sensitive to the scaling of consumption utility. Our formulation also assumes additive separability, though we could extend the model to allow for complementarities in consumption utility to influence behavior similarly to how Kőszegi and Szeidl (2013, footnote 7) suggest extending their focusing model.

how features of options map into consumption dimensions and that we can separately manipulate individual dimensions. It also imposes the norming assumptions *N0* and *N2* — assumptions shared by Kőszegi and Szeidl (2013) — but not our main assumption *N1* that $w(\cdot)$ is decreasing: Indeed, the elicitation can be used to test our assumption against Kőszegi and Szeidl's (2013) that $w(\cdot)$ is increasing. The algorithm elicits consumption utility by examining how a person makes tradeoffs in "balanced choices", for example between $(0,0,a,b,0,0)$ and $(0,0,d,e,0,0)$, where assumptions *N0* and *N2* guarantee that the person will choose to maximize consumption utility. After consumption utilities have been elicited, the algorithm then elicits the weighting function $w(\cdot)$ by examining how ranges in consumption utility influence the rate at which the person trades off utils across dimensions.[9]

The model implies a form of proportional thinking. 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 1. $\tilde{C}$ is *a k-widening of C* if

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

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.

[9] To take a parameterized example, consider

$$w(\Delta) = (1-\rho) + \rho \frac{1}{\Delta + \xi},$$

where $\rho \in [0,1)$ and $\xi \in (0,\infty)$. When $\rho = 0$, the model corresponds to the classical, non-relative-utility model, where a person only considers level differences when making trade-offs. When $\rho > 0$, a marginal change in underlying consumption utility looms smaller when the range is wider. Compared to the underlying utility, people act as if they care less about a dimension the wider the range of utility in that dimension. Note that Assumptions *N2* and *N3* hold: *N3* requires $\rho < 1$ and $\xi > 0$. In the limit case as $\rho \to 1$ and $\xi \to 0$, the actual utility change on a dimension from different choices does not matter, just the percentage change in utility on that dimension.

Proposition 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 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.[10] 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.

To take a more concrete example, Proposition 1 implies that a person's willingness to exert e units of effort to save \$X on a purchase is greater when the relative amount of effort, measured in proportion to the range of effort under consideration, is lower or the relative amount of money saved, measured in proportion to the range of spending under consideration, is higher. In this manner, the model is consistent with evidence used to motivate proportional thinking, such as Tversky and Kahneman's (1981) famous "jacket-calculator" example, based on examples by Savage (1954) and Thaler (1980) and explored further by Azar (2011)—where people are more willing to travel 20 minutes to save \$5 on a \$15 purchase than on a \$125 purchase—so long as not buying is an

[10] 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.

option.[11] Diminishing sensitivity could also explain this pattern, but relative thinking makes the further prediction that traveling 20 minutes to save \$5 on a purchase seems more attractive when it is also possible to travel 50 minutes to save \$11. When not buying at all is an option, the additional savings does not expand the range along the money dimension, but does expand the range of time costs and therefore makes traveling 20 minutes seem small.[12] Proposition 1 is also consistent with van den Assem, van Dolder, and Thaler's (2012) evidence that game show contestants are more willing to cooperate in a variant of the Prisoner's Dilemma when the gains from defecting are much smaller than they could have been: if any non-pecuniary benefits from cooperating are fairly flat in the stakes relative to the monetary benefits from defecting, making the stakes small relative to what they could have been also makes the benefits from defecting appear relatively small.[13]

---

[11]Like explanations based on diminishing sensitivity, ours also relies on the idea that people narrowly bracket spending on a given item. Indeed, Tversky and Kahneman (1981) fix total spending in their example—they compare responses across two groups given the following problem, where one group was shown the values in parantheses and the other was shown the values in brackets:

> Imagine that you are about to purchase a jacket for (\$125) [\$15], and a calculator for (\$15) [\$125]. The calculator salesman informs you that the calculator you wish to buy is on sale for (\$10) [\$120] at the other branch of the store, located 20 minutes drive away. Would you make the trip to the other store?

If people broadly bracketed spending on the two items, then the range of spending would be the same across the two groups, and our model would not be able to account for the difference in the propensity to make the trip.

Our explanation also requires that "don't buy" is in the choice set. While not explicitly stated as an option in the original problem above, we conjecture that people nevertheless contemplate the possibility of not buying the jacket and calculator. We are unaware of examples showing that the jacket-calculator pattern persists in situations where people truly rule out the possibility of not buying, for example if they lose their cell phone and need to buy a new one. To the degree that it does, then the impact is likely coming from diminishing sensitivity.

[12]Contrasting the following example with and without the brackets may also help build intuition: "Imagine you're about to buy a [\$50] calculator for \$25 [by using a 50% off coupon, which was unexpectedly handed to you when you entered the store]. A friend informs you that you can buy the same calculator for \$20 at another store, located 20 minutes drive away. Would you make the trip to the other store?" Range-based relative thinking predicts that people would be less likely to make the trip in the bracketed condition. Diminishing sensitivity can explain this as well but only if we take the reference price to equal the expected price paid rather than zero. However, if we adopt the reference price to equal the expected price paid, diminishing sensitivity can no longer explain the original jacket/calculator pattern.

[13]The contingent-valuation literature discusses related effects in the context of interpreting apparent scope neglect in stated willingness to pay (WTP). For example, Desvouges et al. (1993) asked people to state their willingness to pay to avoid having migratory birds drown in oil ponds, where the number of birds said to die each year was varied across groups. People were completely insensitive to this number in stating their WTP: the mean WTPs for saving 2000, 20,000, or 200,000 birds were \$80, \$78, and \$88, respectively. Frederick and Fischoff (1998) provide a critical analysis of interpretations of such insensitivity—coined the "embedding effect" by Kahneman and Knetsch (1992). While some have been tempted to interpret such evidence as reflecting standard considerations of diminishing marginal rates of substitution, others have argued more plausibly that the insensitivity could reflect range-like effects where, even if the underlying willingness-to-pay function is linear, the displayed willingness-to-pay across subjects could be highly concave. Indeed, studies have found that people are much more sensitive to quantities in within-subject designs. Hsee, Zhang, Lu and Xu (2013) apply a similar principle to develop a method to boost charitable donations. Interestingly, people have argued that within-subject designs may also not accurately elicit true WTP, again because of range-effects. As Frederick and Fischoff (1998, page 116) write:

> we suspect that valuations of any particular quantity would be sensitive to its relative position within

Another property of the model is that the decision-maker's choices will be overly sensitive to the *number* of distinct advantages of one option over another, and insufficiently sensitive to their size. Consider the following consumption vectors

$$c^1 = (2,3,0)$$
$$c^2 = (0,0,5),$$

and assume linear utility. When $C = \{c^1, c^2\}$, then the decision-maker will exhibit a strict preference for c^1 over c^2 despite underlying consumption utility being equal: $2w(2)+3w(3) > 5w(5)$, since $w(\cdot)$ is decreasing by assumption *N1*, implying that $2w(2)+3w(3) > 5w(3) > 5w(5)$.

More starkly, consider a limiting case of assumption *N2* where $\Delta w(\Delta)$ is constant in Δ, so the decision-maker cares only about proportional advantages and disadvantages relative to the range of consumption utility. When comparing two consumption vectors that span the range of consumption utility, each advantage and disadvantage represents 100% of the range and looms equally large. Comparing the two vectors thus reduces to comparing the *number* of advantages and disadvantages in this case. Appendix A.1 develops a more general result on how, all else equal, relative thinking implies that the attractiveness of one consumption vector over another goes up when its advantages are spread over more dimensions or its disadvantages are more integrated.[14]

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 hedonic dimensions.[15] To take an example very similar to one provided by 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

the range selected for valuation but insensitive to *which* range is chosen, resulting in insensitive (or incoherent) values across studies using different quantity ranges.

[14]Appendix A.1 also discusses this result in light of evidence by Thaler (1985) and others on how people have a tendancy to prefer segregated gains and integrated losses, though the evidence on losses is viewed as far less robust.

[15]This feature of the model also means that choice behavior can exhibit cycles. To take an example, suppose utility is linear 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. This same example illustrates that we have to be careful in how we elicit preferences when people are relative thinkers. 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.

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, where the direction of preference depends on 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.[16]

3 Contextual Thinking and Choice-Set Effects

3.1 Classical Choice-Set Effects

The relative-thinker's preference between two alternatives depends on the set of alternatives under consideration. To avoid cumbersome notation, assume each c is measured in utility units throughout this discussion. Consider 2-dimensional consumption vectors, c, c', which have the property that $U(c) = c_1 + c_2 = c'_1 + c'_2 = U(c')$, where $c'_1 > c_1$ and $c'_2 < c_2$. That is, moving from c to c' involves sacrificing some amount on the second dimension to gain some on the first. When c and c' are the only vectors under consideration, the relative-thinker is indifferent between them: $U^N(c'|\{c,c'\}) - U^N(c|\{c,c'\}) = w(c'_1 - c_1)\cdot(c'_1 - c_1) - w(c_2 - c'_2)\cdot(c_2 - c'_2) = 0$. What happens if we add a third consumption vector c''?

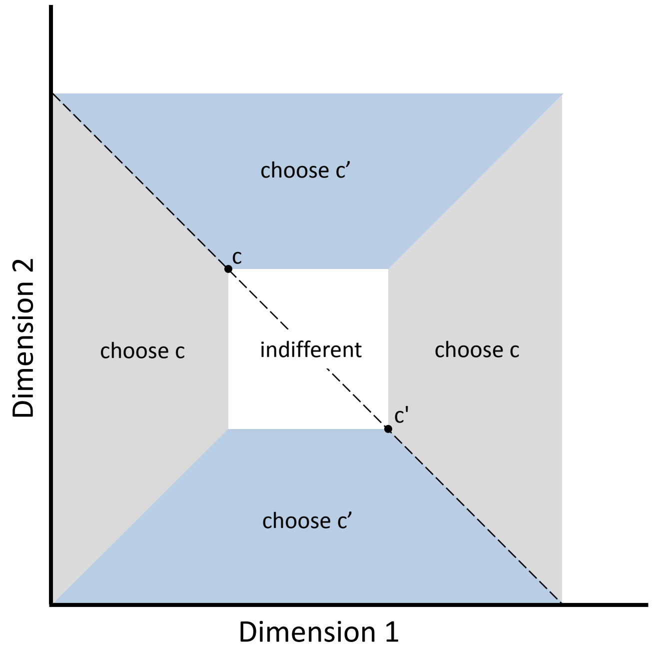
Figure 1 illustrates how the addition of an inferior option influences the preference between c and c'. After presenting the full model that handles choice under uncertainty in Section 4, we will return to the figure and consider the area to the right of the diagonal, which depicts the impact of adding a superior option to the set. Focusing for now on inferior options, when c'' falls in the lighter blue area in the bottom region, its addition expands the range on c''s disadvantageous dimension by more than it expands the range on its advantageous dimension and leads the relative-thinker to choose c' over c. Symmetrically, when c'' falls in the darker grey area in the left region, its addition expands the range on c''s advantageous dimension by more than it expands the range on its disadvantageous dimension and leads the relative-thinker to choose c over c'. Finally, when c'' falls in the white area in the middle region, its addition does not affect the range on either dimension and the relative-thinker remains indifferent between c and c'.

To illustrate, suppose a person is deciding between the following jobs:

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

[16] 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.

Figure 1: The impact of adding the *ex ante* possibility of being able to choose c'' on the relative-thinker's choice from realized choice-set $\{c, c'\}$, where each dimension is measured in utility units.



Job Z. Salary: 120K, Days Off: 119,

where his underlying utility is represented by $U = \text{Salary} + 1000 \times \text{Days Off}$. A relative thinker would be indifferent between jobs X and Y when choosing from $\{X, Y\}$, but instead strictly prefer the higher salary job Y when choosing from $\{X, Y, Z\}$: The addition of Z expands the range of Y's disadvantage relative to X—days off— by more than it expands the range of Y's advantage—salary.[17]

This pattern is consistent with the experimental evidence that adding an inferior "decoy" alternative to a choice set increases subjects' propensity to choose the "closer" of the two initial alternatives—a pattern that Figure 1 illustrates we robustly generate. Famously, experiments have

[17] Our basic results on choice-set effects also highlight a particular way in which the relative thinker expresses a taste for "deals" or "bargains". The addition of the "decoy" job Z makes Job Y look like a better deal in the above example—while getting 10K more salary in moving from Job Y to Job Z requires giving up 70 days off, getting 10K more salary in moving from Job X to Job Y only requires giving up 10 days off. But our model fails to capture some behavioral patterns that may reflect a taste for bargains. Jahedi (2011) finds that subjects are more likely to buy two units of a good at price $\$p$ when they can get one for slightly less. For example, they are more likely to buy two apple pies for \$1.00 when they can buy one for \$.96. While a taste for bargains may undergird this pattern, our model does not predict it: adding one apple pie for \$.96 to a choice set that includes not buying or buying two apple pies for \$1.00 does not expand the range on either the money or the "apple pie" dimensions.

found that adding an alternative that is dominated by one of the initial options, but not the other, increases the preference for the induced "asymetrically dominant" alternative.[18] This effect, called the *asymmetric dominance effect* or *attraction effect* was initially shown by Huber, Payne and Puto (1982), and has been demonstrated when subjects trade off price vs. quality or multiple quality attributes of consumer items (e.g., Simonson 1989), the probability vs. magnitude of lottery gains (e.g., Soltani, De Martino, and Camerer 2012), and various other dimensions including demonstrations by Herne (1997) over hypothetical policy choices and Highhouse (1996) in hiring decisions.[19] Consistent with our model, similar effects (e.g., Huber and Puto 1983) are found when the decoy is not dominated but "relatively inferior" to one of the two initial alternatives. There is also evidence that context effects are more pronounced when the decoy is positioned "further" from the original set of alternatives, increasingly expanding the range of one of the dimensions (Heath and Chatterjee 1995; Soltani, De Martino, and Camerer 2012).[20]

Our basic results on choice-set effects are not shared by Bordalo, Gennaioli, and Shleifer (2013) or other recent models, including Kőszegi and Szeidl (2013) and Cunningham (2013), that likewise model such effects as arising from features of the choice context influencing how attributes of different options are weighed. We describe this in detail in Appendix C.[21] Here, we primarily confine attention to comparing our predictions to Kőszegi and Szeidl's (2013) because their model also assumes that decision weights are solely a function of the range of consumption utility, which makes it the simplest to compare. Since it makes the opposite assumption on how the range matters, namely that decision weights are *increasing* in the range, it makes opposite predictions to ours in all two-dimensional examples along the lines illustrated in Figure 1. Figure 2 illustrates Kőszegi

[18]We emphasize laboratory evidence on attraction effects because we believe it directly speaks to basic predictions of choice-set dependent models. For various reasons, some of which are emphasized below in Sections 3.2 and 7, we are less convinced that these effects are necessarily important in the field. Recent studies that provide other reasons to question the practical significance of attraction effects include Frederick, Lee, and Baskin (2014) as well as Yang and Lynn (2014).

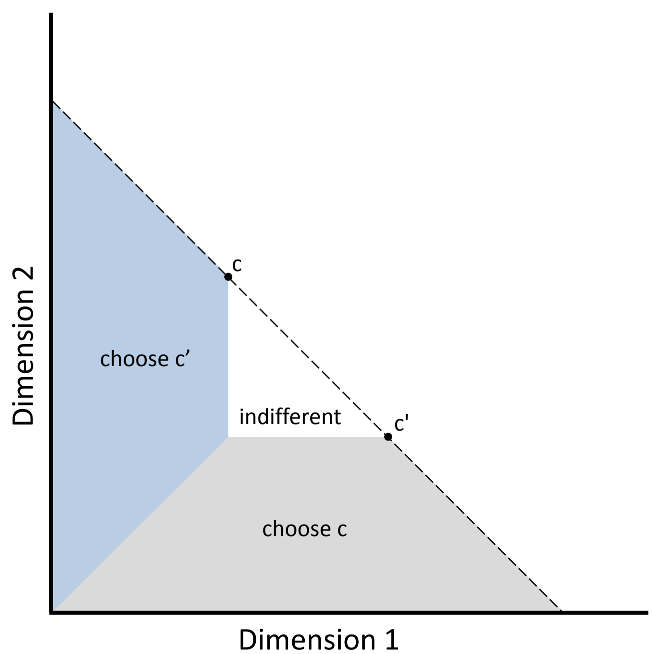
[19]While initial demonstrations by Huber, Payne, and Puto (1982) and others involved hypothetical questionnaires, context effects like asymmetric dominance have been replicated involving real stakes (Simonson and Tversky 1992, Doyle, O'Connor, Reynold, and Bottomley 1999, Herne 1999, Soltani, De Martino, and Camerer 2012). They have also been demonstrated in paradigms where attempts are made to control for rational inference from contextual cues (Simonson and Tversky 1992, Prelec, Wernerfelt, and Zettelmeyer 1997, Jahedi 2011)—a potential mechanism fomalized by Wernerfelt (1995) and Kamenica (2008). Closely related is the "compromise effect" (Simonson 1989), or the finding that people tend to choose middle options.

[20]Huber, Payne, and Puto (1982) suggest a mechanism similar to ours to account for attraction effects and asymmetric dominance, though Huber and Payne (1983) argue that it is difficult to explain the evidence as resulting from "range effects" because the likelihood of choice reversals seems to be insensitive to the magnitude of the range induced by the position of the decoy. Wedell (1991) shows something similar, and Simonson and Tversky (1992) cite all this evidence as "rejecting" the range hypothesis. However, meta-analysis by Heath and Chaterjee (1995) suggest that range effects do in fact exist in this context, as have more recent studies, such as Soltani, De Martino, and Camerer (2012).

[21]Soltani, De Martino, and Camerer (2012) present a model that shares similar motivations, but is written with a different focus: Their model shows how the biophysical limits of neural representations can account for range effects in a specific choice context. Our model instead takes range effects as given, but fleshes out the assumptions necessary to broaden the domain of application to a greater variety of economic contexts.

and Szeidl's (2013) predictions on how adding inferior option c'' to $\{c, c'\}$ influences the person's choice between c and c' when initially indifferent, which can be compared to our predictions illustrated in the area to the left of the diagonal line in Figure 1.Their model says that adding a more extreme, inferior, option to the choice set leads the person to prefer the *further* of the two superior options, since the addition expands the range of the closer option's disadvantage relative to the other superior alternative by more than it expands the range of its advantage and hence attracts attention to its disadvantage. Their model says that adding Job Z will lead people to choose Job X because its addition draws attention to Days Off. The predictions of their model in two-dimensional examples seem hard to reconcile with the laboratory evidence on attraction effects summarized above. One possibility we explore in Section 6 is that "focusing effects", along the lines that Kőszegi and Szeidl (2013) model, may be more important in choice problems involving many dimensions than in two-dimensional problems like these.

Figure 2: Kőszegi and Szeidl's (2013) predictions on the impact of adding an inferior c'' to the comparison set on the person's choice between c and c', assuming each option is measured in utility units.



None of these models, including ours, capture certain forms of the compromise effect (Simonson 1989; Tversky and Simonson 1993). In our model, a person who is indifferent between 2-dimensional options c, c', and c'' without relative thinking will remain indifferent with relative thinking: he will not display a strict preference for the middle option. Likewise, Bordalo, Gennaioli, and Shleifer (2013) observe that their model does not mechanically generate a preference for choosing "middle" options. Kőszegi and Szeidl (2013) and Cunningham (2013) also cannot generate these effects.[22]

An alternative interpretation for why trade-offs depend on ranges along consumption dimensions—giving rise to the sorts of choice-set effects we emphasize in this section—is that this follows as a consequence of inference from contextual cues, broadly in the spirit of mechanisms proposed by Wernerfelt (1995) and Kamenica (2008). In some circumstances, a person who is uncertain how to value an attribute dimension may rationally place less weight on that dimension when its range is wider, perhaps by guessing that hedonic ranges tend to be similar across dimensions, and therefore guessing that the hedonic interpretation for a change in a dimension is inversely related to the range in that unit. While we believe that such inference mechanisms likely play an important role in some situations, evidence suggests that they do not tell a very full story. There is evidence of range effects in trade-offs involving money and other dimensions that are easily evaluated, such as in Soltani, De Martino, and Camerer (2012), where people make choices between lotteries that vary in the probability and magnitude of gains. Mellers and Cooke (1994) show that range effects are found even when attributes have a natural range that is independent of the choice set, for example when they represent percentage scores, which naturally vary between 0 and 100.

3.2 The Limits of Choice-Set Effects

The first portion of this section showed that relative thinking implies some classical choice-set effects. We now provide bounds for choice-set effects that result from relative thinking. Given any two options c and c', the following proposition supplies necessary and sufficient conditions on their relationship for there to exist a choice set under which c' is chosen over c.

Proposition 2.

1. *Assume that each $u_k(c_k)$ is unbounded above and below. For $c, c' \in \mathbb{R}^K$ with $U(c') \geq U(c)$, either c' would be chosen from $\{c, c'\}$ or there exists c'' that is undominated in $\{c, c', c''\}$ such that c' would be chosen from $\{c, c', c''\}$.*

2. *Assume that each $u_k(c_k)$ is unbounded below. For $c, c' \in \mathbb{R}^K$ with $U(c) \neq U(c')$ there is a C*

[22] For a review of models aimed at capturing the compromise effect, see Kivetz, Netzer, and Srinivasan (2004). None of these models capture the array of results that we emphasize.

containing $\{c, c'\}$ such that c' is chosen from C if and only if

$$\sum_{i \in A(c',c)} w(\delta_i(c',c)) \cdot \delta_i(c',c) + \sum_{i \in D(c',c)} w(\infty) \cdot \delta_i(c',c) > 0, \qquad (1)$$

where $A(c',c) = \{k : u_k(c'_k) > u_k(c_k)\}$ denotes the set of c''s advantageous dimensions relative to c and $D(c',c) = \{k : u_k(c'_k) < u_k(c_k)\}$ denotes the set of c''s disadvantageous dimensions relative to c.

Part 1 of Proposition 2 shows that if c' yields a higher un-normed utility than c, then there exists some choice set where it is chosen over c. This part only relies on *N0(d)*, or in particular that the person makes a utility-maximizing choice from C whenever the range of utility on each dimension is the same given C, or whenever $\Delta_j(C)$ is constant in j.[23] The intuition is simple: so long as utility is unbounded, one can always add an option to equate the ranges across dimensions. For example, while we saw before that the relative thinker prefers $(2,3,0)$ over $(0,0,5)$ from a binary choice set, this finding says that because the un-normed utilities of the two options are equal there exists a choice set containing those options under which the relative thinker would instead choose $(0,0,5)$. In particular, a person would always choose $(0,0,5)$ from $\{(2,3,0),(0,0,5),(5,-2,0)\}$.

The second part of the proposition uses the additional structure of Assumptions *N1-N2* to supply a necessary and sufficient condition for there to exist a comparison set containing $\{c, c'\}$ such that the person chooses c' over c.[24] For intuition on where condition (1) comes from, it is equivalent to asking whether c' would be chosen over c when the comparison set is such that the range over its advantageous dimensions are the smallest possible (i.e., $u_i(c'_i) - u_i(c_i)$), while the range over its disadvantageous dimensions are the largest possible (i.e., ∞). In the classical model (with a constant w_k), this condition reduces to $U(c') > U(c)$. In the limiting case—ruled out by *N3* – where $w(\infty) = 0$, the condition is that c' has *some* advantageous dimension relative to c (i.e., is not dominated). More generally, as spelled out in Appendix A.2, the difference in un-normed utilities between the options cannot favor c "too much" and c' must have some advantages relative to c that can be magnified.[25]

Taken together, the two parts of the proposition say that the impact of the comparison set is

[23] As a result, the first part of Proposition 2 also holds for Kőszegi and Szeidl (2013).

[24] As the proof of Proposition 2 makes clear, the conclusions are unchanged if C is restricted such that each $c'' \in C \setminus \{c, c'\}$ is undominated in C.

[25] These comparative statics can perhaps be seen more clearly by re-writing inequality (1) under *N3* as

$$U(c') - U(c) + \sum_{i \in A(c',c)} \left(\frac{w(\delta_i(c',c))}{w(\infty)} - 1 \right) \cdot \delta_i(c',c) > 0,$$

which highlights how the inequality depends on the difference in un-normed utilities as well as whether c' has advantages relative to c.

bounded in our model. In particular, if $U(c') > U(c)$, it is always possible to find a comparison set under which the agent displays a preference for c' over c. However, it is only possible to find a comparison set under which the agent displays a preference for c over c' when (1) holds.

Our model implies two additional limits to comparison-set effects. First, our model says that people maximize un-normed consumption utility when decisions involve sufficiently large stakes: Supposing each $u_k(\cdot)$ is unbounded, one can show that for all comparison sets C, there exists a $\bar{t} > 0$ such that if c' is an un-normed utility-maximizing choice from C, then $t \cdot c'$ is a normed utility-maximizing choice from $t \cdot C$ for all $t > \bar{t}$. One intuition is that absolute differences scale up with t, but proportional differences do not, so absolute differences dominate decision-making as t gets large.

Second, as shown in Appendix A.2 (Proposition 8), for any option c, there exists a choice set C containing c together with "prophylactic decoys" such that c will be chosen and, for any expansion of that set, only options that yield "approximately equivalent" un-normed utility to c or better can be chosen. This means, roughly, that once C is available no decoys can be used to leverage range effects to make a consumer choose any option inferior to c. With unbounded utility and Assumption *N3*, 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. One potential application of this result lies in examining competition in a product market. For example, a firm that wishes to sell some target product can always market other products that would not be chosen, but prevent other firms from introducing options that frame sufficiently inferior products as superior. This suggests that relative thinking may influence the options that are offered to market participants by more than it influences ultimate choices.[26]

4 The Full Model

We now present the full model which allows for uncertainty. The decision-maker chooses between lotteries on $\mathbb{R}^K$, and the choice set is some $\mathscr{F} \subset \Delta(\mathbb{R}^K)$. This captures standard situations where a decision-maker chooses between monetary risks, but also situations where a decision-maker makes plans prior to knowing the exact choice set he will face. For example, if the decision-maker makes plans knowing he faces choice set $\{c, c'\}$ with probability q and choice set $\{c, c''\}$ with probability $1-q$, his choice is between lotteries in $\mathscr{F} = \{(1,c), (q,c';1-q,c), (q,c;1-q,c''), (q,c';1-$

[26] These conclusions depend on the market structure as well as the technologies that are available to firms. Monopolists may have an incentive to market decoys that get consumers to buy inferior products. In competitive situations, consumers may still buy inferior products if prophylactic decoys are prohibitively costly to market, for example if they must be built in order to market them or if there is sufficient consumer heterogeneity that some consumers will actually demand these products if they are offered. A more complete analysis of competition with decoys, which is left for future work, would need to grapple with such issues.

$q, c'')\}$. Since the earlier definition of the range of consumption utility along a dimension only applies to riskless choice, we need to extend the definition to address such situations.

Perhaps the simplest formulation would take the range along a dimension to equal the range induced by $\tilde{C} = \{c \in \mathbb{R}^K | c$ is in the support of some $F \in \mathscr{F}\}$, in which case the range in the above example would equal $\Delta_k = \max_{\tilde{c} \in \{c,c',c''\}} u_k(\tilde{c}_k) - \min_{\tilde{c} \in \{c,c',c''\}} u_k(\tilde{c}_k)$. A problem with this formulation is that it treats low and high probability outcomes the same: The range along a dimension is the same whether $q \approx 1$ and the person knows with near certainty that c'' will not be an option and when $q \approx 0$ and the person knows with near certainty that c' will not be an option. To take another example, it would not capture an intuition that \$10 feels like a lower return to effort when \$15 was not just possible, but probable.

Another option would be to take the range along a dimension to equal that induced by $\tilde{C} = \{c \in \mathbb{R}^K | c = E[F]$ for some $F \in \mathscr{F}\}$, where $E[F] = \int c \cdot dF(c)$ is the expected value of c under F. In this case, the range along a dimension in the above example would be the range implied by $\tilde{C} = \{c, (1-q) \cdot c + q \cdot c', q \cdot c + (1-q) \cdot c'', q \cdot c' + (1-q) \cdot c''\}$. While this formulation would successfully treat high and low probability outcomes differently, it has the issue that only the range of expected values *across* lotteries would determine the range used to norm outcomes, and not the range of outcomes in the support of *given* lotteries. This formulation would say that a person would norm outcomes the same way when choosing between lotteries $\{(1/2, -100; 1/2, 110), (1/2, -10; 1/2, 15)\}$ as when choosing between $\{(1/2, -20; 1/2, 30), (1/2, -10; 1/2, 15)\}$ since $1/2(-100) + 1/2(110) = 1/2(-20) + 1/2(30)$; our intuition is that he would instead be less sensitive to a fixed difference in the former case.

To construct the range along a dimension in a way that both depends on probabilities as well as within-lottery ranges, we summarize every lottery's marginal distribution over $u_k(c_k)$ by the mean plus or minus its variation around the mean, where variation is measured by something akin to the standard deviation around the mean, and then take the range along a dimension to equal the difference between the maximal and minimal elements across the summarized distributions. The actual measure of variation we use is the L-scale or 1/2 its mean difference, also known as 1/2 the "average self-distance" of a lottery. We believe very little of our analysis would qualitatively change if classical standard deviation or other notions of dispersion were used.[27]

Formally, given a comparison set $\mathscr{F}$, we define the range along dimension k to equal

$$\Delta_k(\mathscr{F}) = \max_{F \in \mathscr{F}} \left(E_F[u_k(c_k)] + \frac{1}{2} S_F[u_k(c_k)] \right) - \min_{F \in \mathscr{F}} \left(E_F[u_k(c_k)] - \frac{1}{2} S_F[u_k(c_k)] \right), \tag{2}$$

[27]Yitzhaki (1982) argues that the mean difference provides a more central notion of statistical dispersion in some models of risk preference since it can be combined with the mean to construct necessary conditions for second-order stochastic dominance. Kőszegi and Rabin (2007) show it to be especially relevant for models of reference-dependent utility, which also means that this definition may facilitate combining our model with reference dependence.

where $E_F[u_k(c_k)] = \int u_k(c_k)dF(c)$ equals the expectation of $u_k(c_k)$ under F, and $S_F[u_k(c_k)] = \int\int |u_k(c'_k) - u_k(c_k)|dF(c')dF(c)$ is the "average self-distance" of $u_k(c_k)$ under F, or the average distance between two independent draws from the distribution. Note that the range along a dimension collapses to the previous riskless specification when all lotteries in $\mathscr{F}$ are degenerate. Note also that when $K = 1$ and $\mathscr{F} = \{F\}$ is a singleton choice set, then we have $\Delta(\mathscr{F}) = S_F[u(c)]$.[28]

We assume the weights w_k satisfy the following assumptions, which generalize the conditions from Section 2.

Norming Assumptions:

N0. The weights w_k are given by $w_k = w(\Delta_k(\mathscr{F}))$, where $\Delta_k(\mathscr{F})$ is given by (2).

N1. $w(\Delta)$ is a differentiable, decreasing function on $(0,\infty)$.

N2. $w(\Delta)\cdot\Delta$ is defined on $[0,\infty)$ and is strictly increasing.

Assumption *N0* expands the definition of the range from *N0(d)* to the more general definition of (2), while the rest of the norming assumptions remain as they were in the deterministic case. Given the comparison set, the decision-maker evaluates probabability measure F over $\mathbb{R}^K$ according to:

$$U^N(F|\mathscr{F}) = \int U^N(c|\mathscr{F})dF(c),$$

where $U^N(c|\mathscr{F}) = \sum_k w(\Delta_k(\mathscr{F}))\cdot u_k(c_k)$ as in the riskless case.[29]

To build more intuition, consider the following examples.

Example 1.

$$\mathscr{F} = \{(1/2,-100;1/2,110),(1/2,-10;1/2,15)\}$$
$$\mathscr{F}' = \{(1/2,-20;1/2,30),(1/2,-10;1/2,15)\}.$$

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

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

[29] Some experimental evidence on attraction effects when people choose between lotteries of the form "gain g with probability p" suggests a different approach to applying the model to choice under uncertainty: treat (Probability Gain, Magnitude Gain) as attributes. For example, evidence from Soltani, De Martino and Camerer (2012) indicates that adding gamble (Probability Gain, Magnitude Gain) $= (.6,15)$ to $\{(.3,50),(.7,20)\}$ pushes subjects towards $(.7,20)$. However, we suspect that these experimental results are highly sensitive to precisely how lotteries are framed as vectors of attributes. For example, we conjecture that we would see a different pattern of results if lotteries are framed in terms of (Standard Deviation, Expected Value) instead of (Probability Gain, Magnitude Gain). Such a change can alter whether options are viewed as asymmetrically dominant, for example the (.7, 20) lottery from above does not dominate (.6, 15) in (Standard Deviation, Expected Value) space as it has a greater standard deviation. We believe that the results we highlight are likely more robust to framing and elicitation techniques than results that would come from mechanically applying the model using something like (Probability Gain, Magnitude Gain) as attributes.

In this case, the range given $\mathscr{F}$ is $\Delta(\mathscr{F}) = 105$ and the range given $\mathscr{F}'$ is $\Delta(\mathscr{F}') = 25$. The range is larger in the case of $\mathscr{F}$ than $\mathscr{F}'$, illustrating that the range of outcomes in the support of given lotteries matters, not just the range of average outcomes across lotteries. □

Example 2. The decision-maker makes plans knowing he faces choice set $\{c, c'\}$ with probability q and choice set $\{c, c''\}$ with probability $1-q$, so $\mathscr{F} = \{(1,c),(q,c';1-q,c),(q,c;1-q,c''),(q,c';1-q,c'')\}$. Supposing $c = (0,0), c' = (-1,1)$, and $c'' = (-1,2)$, the range along the first dimension is $\Delta_1(\mathscr{F}) = 1$ and the range along the second is $\Delta_2(\mathscr{F}) = 2-q^2 \in [1,2]$. Importantly, the range along the second dimension is decreasing in q and tends towards 1 (resp. 2) as q tends towards 1 (resp. 0), illustrating both a) that high and low probability outcomes are treated differently (and in the intuitive direction) in our formulation and b) that, as uncertainty about the choice set vanishes, ranges converge to those associated with the realized choice set. □

The second example shows that when the choice set people ultimately face is uncertain, they may be influenced by options that are not in the realized choice set. Return to Figure 1 and consider what happens when the comparison set can differ from the realized choice set, which will allow us to analyze how the relative-thinker's choice between c and c' is influenced by having contemplated the possibility of being able to choose a *superior* option.[30] For concreteness, consider the situation where the person makes plans prior to knowing which precise choice set he will face and makes choices from $\{c, c'\}$ with probability $1-q$ and makes choices from $\{c, c', c''\}$ with probability q—under this interpretation, Figure 1 illustrates the relative-thinker's choice when $\{c, c'\}$ is realized and $q > 0$. The area to the right of the diagonal line in Figure 1 illustrates how adding a superior option to the comparison set influences the relative-thinker's preference between c and c'. In the grey area to the right of the line (the blue area can be symmetrically analyzed), the addition of c'' to the comparison set expands the range of c''s advantageous dimension by more than it expands the range of its disadvantageous dimension, pushing the relative-thinker to choose c from $\{c, c'\}$.

To illustrate, consider a student deciding whether to take an odd job that pays \$100 with commensurate cost in terms of foregone leisure, so we can think of him as choosing from the choice set $\{(100,-100),(0,0)\}$. The analysis says that the student is less likely to choose to work if his comparison set includes a similar job that pays \$125, because this expands the range on the money dimension, and makes \$100 seem like a lower return to effort.

These results connect with a smaller experimental literature that examines how making subjects aware of a third "decoy" alternative—that could have been part of the choice set but is not—influences their preferences between the two alternatives in their realized choice set. While the overall evidence seems mixed and debated, as far as we are aware the cleanest experiments from

[30] Some experimentalists (e.g., Soltani, De Martino and Camerer 2012) induce such a wedge between choice and comparison sets by having subjects first evaluate a set of alternatives during an "evaluation period" and then quickly make a selection from a random subset of those alternatives during a "selection period".

the perspective of our model—for example, that make an effort to control for rational inference from contextual cues—have found that including "asymmetrically dominant" (or "close to dominant") decoys to the comparison set decreases experimental subjects' propensity to choose the asymmetrically dominated target when the decoy is not present in the choice set (e.g, Soltani, De Martino, and Camerer 2012). Relatedly, Jahedi (2011) finds that subjects are less likely to buy a good (for example, an apple pie) if they are aware that there was some probability they could have bought two of the same good for roughly the same price (for example, that there was some probability of getting a two-for-one deal on apple pies).[31]

5 Extended Economic Implications

This section develops some applications of the model. We first consider how uncertainty influences the rate at which people trade off utility across dimensions, and go on to study choice over time.

5.1 Tradeoffs Across Dimensions Under Uncertainty

Consider a simple example where somebody chooses how much effort to put into a money-earning activity, and has two consumption dimensions—money and effort. But he also earns money from another stochastic source. His un-normed utility is given by

$$U = r \cdot e \pm k - f \cdot e,$$

where $e \in \{0, 1\}$ denotes his level of effort, r equals the return to effort, f represents the cost to effort, and $\pm k$ indicates an independent 50/50 win-k/lose-k lottery.

A person maximizing expected utility would choose $e^* = 1$ if and only if $r/f \geq 1$. Notably, his effort is independent of k.

By contrast, our model says that increasing k will decrease effort: the more income varies, the smaller will seem an additional dollar of income from effort, and so the less worthwhile will be the effort. To see this, note that (given the formula for ranges in stochastic settings outlined in Section 4) the range in consumption utility along the income dimension is $r + k$, while the range along the

[31] A contrasting finding in the experimental literature has come to be labeled the "phantom decoy effect" (Pratkanis and Farquhar 1992): presenting a dominant option declared to be unavailable can bias choice towards the similar dominated option (Highhouse 1996; Pettibone and Wedell 2007; Tsetsos, Usher and Chater 2010). Some of this may be due to rational inference: taking an example inspired by Highhouse (1996), if job candidates vary in interview ability and test scores, then knowledge that a job candidate with extremely high interview ability and medium test scores is no longer on the market may provide a signal that interview ability is more important than test scores, leading a person to select a candidate with high interview ability and medium test scores over a candidate with medium interview ability and high test scores.

effort dimension is f. Normed utility then equals

$$U^N = w(r+k)\cdot(r\cdot e \pm k) - w(f)\cdot f\cdot e,$$

so expected normed utility equals $EU^N = w(r+k)\cdot r\cdot e - w(f)\cdot f\cdot e$. The person then works so long as

$$\frac{r}{f} \geq \frac{w(f)}{w(r+k)},$$

where the right-hand side of this inequality is increasing in k, and greater than 1 for large enough k. This implies that, in contrast to the expected-utility maximizer, the relative thinker is less likely to work when k is larger: increasing uncertainty on a dimension decreases his sensitivity to incremental changes in utility along that dimension.[32]

To state a more general result, for lotteries $H, H' \in \Delta(\mathbb{R})$, let $H+H'$ denote the distribution of the sum of independent draws from the distributions H and H'. Additionally, for lotteries $F_i \in \Delta(\mathbb{R})$, $i = 1,\ldots,K$, let $(F_1,\ldots,F_K) \in \Delta(\mathbb{R}^K)$ denote the lottery where each c_i is independently drawn from F_i.

Proposition 3. *Assume $K = 2$ and each $u_i(\cdot)$ is linear for $i = 1,2$.*

1. *For $F_1, F_2 \in \Delta(\mathbb{R})$ and $G_1, G_2 \in \Delta(\mathbb{R}^+)$, if (F_1,F_2) is chosen from $\{(F_1,F_2),(F_1-G_1,F_2+G_2)\}$, then (F_1,F_2') is chosen from $\{(F_1,F_2'),(F_1-G_1,F_2'+G_2')\}$ whenever F_2' is a mean-preserving spread of F_2 and G_2' is a mean-preserving spread of G_2. Moreover, the choice is unique whenever $F_2' \neq F_2$ or $G_2' \neq G_2$.*

2. *For $F_1, F_2 \in \Delta(\mathbb{R}^+)$, suppose the person faces the distribution over choice sets of the form $\{(0,0),(-\tilde{x},\tilde{y})\}$ that is induced by drawing $\tilde{x}$ from F_1 and $\tilde{y}$ from F_2. If $(0,0)$ is preferred to realization $(-x,y)$ given the resulting comparison set, then $(0,0)$ is strictly preferred to*

[32] If marginal utility over money is convex, as is often assumed in explaining precautionary savings using the neoclassical framework, then income uncertainty will have the opposite effect. In such a case, expected marginal utility is increasing in income uncertainty, so higher income uncertainty will increase the propensity to exert effort to boost income. If the less conventional assumption of concave marginal utility is made, then more uncertainty would, as in our model, decrease the value on money. But in either case, the effects would be calibrationally small for modest increases in uncertainty. Loss aversion, by constrast, makes a bigger and less ambiguous opposite prediction to ours. Consider a situation in which people are able to commit in advance to take effort. Applying the concept of choice-acclimating equilibria from Kőszegi and Rabin (2007) and assuming linear consumption utility, loss aversion predicts no impact of higher income uncertainty on the propensity to exert effort: The decision is determined solely by consumption utility. Consider a second situation in which the opportunity to exert effort comes as a surprise and the person thus previously expected to not exert effort. In this case, loss aversion predicts that the presence of a 50/50 win-k/lose-k lottery over money increases the person's willingness to exert effort: Returns to effort are shifted from being assessed as increasing gains to partially being assessed as reducing losses.

$(-x,y)$ if instead the distribution over choice sets is induced by drawing $\tilde{x}$ from F_1 and $\tilde{y}$ from $F_2' \neq F_2$, where F_2' first order stochastically dominates a mean-preserving spread of F_2.

Part 1 of Proposition 3 generalizes the above example, and says that if the person is unwilling to sacrifice a given amount from one dimension to the other, he will not do so if the second dimension is made riskier. Notably, the proposition extends the example by allowing the person to influence the amount of risk he takes. To illustrate, consider a simple modification of the example where exposure to risk goes up in effort, and utility equals $U = e \cdot (r \pm k) - f \cdot e$. The proposition says that, again, the worker is less likely to exert effort when the amount of income uncertainty, k, is higher: the wider range in the monetary dimension reduces the worker's sensitivity to incremental changes in money.

Part 2 says that a person becomes less willing to transfer a given amount of utility from one dimension to a second when the background distribution of potential benefits on the second dimension becomes more dispersed or shifted upwards. For example, if a person is indifferent between exerting effort e to gain \$100 if he made plans knowing \$100 is the return to effort, he will not exert effort if, *ex ante*, he placed equal probability on earning \$50, \$100, or \$150. And such a person will be even less likely to exert effort for \$100 if he were *ex ante* almost sure to be paid \$150 for effort, since this further expands the range on the money dimension and makes earning \$100 feel even smaller.

Proposition 3 applies to tradeoffs beyond those involving effort and money. For example, it could be that $U = c_1 + c_2$, where $c_1 \in \{0,1\}$ represents whether the person has a good, such as shoes, and $c_2 \in \mathbb{R}$ represents dollar wealth. In this case, Part 2 says that the person will be more likely to buy when prices are more uncertain *ex ante*. For example, if a person is indifferent between buying and not buying at price \$20 if he knew that \$20 would be the price, then he strictly prefers to buy at \$20 if, *ex ante*, he placed equal probability on the price being \$15, \$20, or \$25. Part 2 also says that the person will be more likely to buy at a fixed price when he expected higher prices. In other words, people's reservation prices will go up in expected prices, which may shed light on why retailers can benefit from advertising high list prices for goods they are trying to sell: making goods occassionally available at inflated list prices makes the regular "discounted" price feel smaller.[33]

Uncertainty not only makes the person less willing to transfer utility across dimensions, but also attenuates his response to incentives. To see this, enrich the above example by allowing the

[33] Bordalo Gennaioli and Shleifer (2013) offer a related psychological account based on the idea that advertising high prices can draw people's attention away from paying the regular price by bringing it closer to average in the set of prices under consideration. This explanation does not seem to mesh with the fact that sellers loudly advertise discounts, which presumably draws buyers' attention to prices. In our model, sellers are motivated to draw people's attention to the existence of high prices to make sales prices feel smaller. A difference in predictions is that Bordalo, Gennaioli, and Shleifer's explanation also says that making the list price very high would actually backfire by drawing attention to price by pushing the regular price further from the average.

person's effort choices to be continuous. The person chooses the amount of effort $e \in [0,1]$ to engage in a project at cost $1/2 \cdot e^2$, where e equals the probability that a project will be successful. If successful, the project yields return $r+y$, where $r > 0$ and y is a mean-zero random variable drawn according to distribution F with $y > -r$ for all y in the support of F. The person knows y prior to his choice of effort, but makes contingent plans only given knowledge of F.

The range on the money dimension is $r + 1/2 \cdot S(F)$ and the range on the effort dimension is $1/2$, so for each y the person chooses an amount of effort to solve

$$\max_{e \in [0,1]} w(r + 1/2 \cdot S(F)) \cdot e \cdot (r+y) - w(1/2) \cdot 1/2 \cdot e^2.$$

Taking first-order conditions, we solve for contingent effort choice

$$e(y) = \frac{w(r + 1/2 \cdot S(F))}{w(1/2)} \cdot (r+y). \tag{3}$$

As before, we see that the person works less hard when the *ex ante* range on returns is larger: $\partial e / \partial S(F) < 0$. The novel comparative static is that the person's effort choices are also less sensitive to the realized return when this range is larger: $\partial^2 e / \partial S(F) \partial (r+y) < 0$. For example, the effort choices of a person who places a 50/50 chance on getting paid an \$8 or \$10 piece-rate for a task will appear more sensitive to the realized piece rate than a person who initially places equal probability on \$7, \$8, \$9, \$10, and \$11.

Summarizing, in the context of effort decisions, the model makes several novel predictions in uncertain environments. First, a person is less likely to exert effort for a fixed return when *ex ante* or *ex post* income uncertainty is greater. Second, the same is true when the person expected to earn more when forming plans. Third, a person's effort choices are less sensitive to the realized return when he faced greater uncertainty in *ex ante* possible returns.

These predictions of our model contrast with those of other choice-set dependant or relative-thinking models with which we are familiar. While we know of no strong evidence that speaks to our predictions, field and laboratory findings provide suggestive supporting evidence. There is evidence, for instance, that insuring farmers against adverse weather shocks such as drought can increase their willingness to make high-marginal-return investments.[34] While this investment response could in principle result from risk aversion if investment returns negatively covary with the marginal utility of consumption—for example, investment in fertilizer could have a lower return when rainfall and consumption are low and the marginal utility of consumption is then high—it may also in part result from similar mechanisms to those we highlight: Our model predicts a

[34]See, e.g., Karlan, Osei, Osei-Akoto, and Udry (2013), Cole, Gine, and Vickery (2011), and Mobarak and Rosenzweig (2012), who find that uncertainty reduction through insurance seems to increase investments such as fertilizer use and weeding.

positive investment response *even when investment returns are uncorrelated with the newly insured risk*, thus broadening the set of circumstances where we would expect expanding insurance coverage to boost profitable investment.

Suggestive evidence also arises from laboratory findings on relative pay and labor supply. Bracha and Gneezy (2012) find that the willingness to complete a task for a given wage is inversely related to previous wages offered for a related task. People are less likely to show up to complete a survey for either \$5 *or* \$15 if they were previously offered \$15 to complete a related survey than if they were previously offered \$5.[35] Although a cleaner test of our model's predictions would more directly manipulate expectations, this is consistent with the model if, as seems plausible, expectations of future wages are increasing in the size of previous wages: Increasing the probability attached to a higher wage offer increases the range attached to money, thereby increasing the reservation wage.

5.2 Tradeoffs Across Time Under Uncertainty

We now demonstrate how relative thinking may induce impatient-seeming behavior when a person trades off consumption across periods of time. The reason is that the utility of spending an additional amount, say \$10, represents a greater proportion of the range of today's utility than it does relative to the range of future utility. This is for at least two reasons: uncertainty will impact future utility more than current utility, and—given concave utility—the utility of being able to spend some fixed amount in the future is greater than the utility of spending that amount today. "Range-based present bias" generates comparative-static predictions not implied by quasi-hyperbolic discounting (Strotz 1955; Laibson 1997). Such present bias increases with greater uncertainty in future consumption utility and with longer horizons of consumption. Under conditions we discuss in this section, the model also says a person will act less present-biased if he is frequently allotted a small amount to consume than if he is infrequently allotted a large amount, and on later days in a pay period than on paydays. The overspending induced by range-based present bias can even be on goods that are not immediately consumed: a person is more likely to buy a good if he can buy it on a payday, even if he has to wait to consume the good.

Suppose a person has I dollars to allocate across $(c_1, \ldots, c_T)$, where each c_t represents consumption in period t. To simplify, we assume the person does not discount and the interest rate is zero. In expectation, the person's consumption utility is $\sum_t u(c_t)$, where $u(0) = 0, u'(\cdot) > 0$, and $u''(\cdot) < 0$. A person who is not a relative thinker will then consume $c_t = I/T$ for all t.

We begin by considering the case where $T = 2$ and second-period utility is uncertain. While

[35] And, as one would expect and as embedded in our model, people like high wages: People are more likely to show up to complete a survey for \$15 if they were previously offered \$15 than to show up to complete a survey for \$5 if they were previously offered \$5.

utility in the first period is given by $u(c_1)$, we assume utility in the second is given by $u(c_2) \pm k$ where $\pm k$ denotes a 50/50 gain k, lose k lottery. Uncertainty in second-period utility can be thought of as arising from exogeneous shifts to environmental factors (e.g., the weather) or health shocks, for example. Because the person has I to spend across the two periods, the most he can consume in the first period is I and the range of first-period consumption utility equals $\Delta_{t=1} \equiv u(I) - u(0) = u(I)$. Using the formula for the range under uncertainty, i.e., the formula for $\Delta_j(\mathscr{F})$, the range of second-period consumption utility equals $\Delta_{t=2} \equiv u(I) + k$.[36]

The relative thinker then chooses c_1 to solve

$$\max_{c_1 \in [0,I]} w(\Delta_{t=1}) \cdot u(c_1) + w(\Delta_{t=2}) \cdot u(I - c_1),$$

so he chooses c_t as if he were a quasi-hyperbolic (β, δ) discounter (O'Donoghue and Rabin 1999) with $\delta = 1$ and "effective discount factor"

$$\tilde{\beta} = w(\Delta_{t=2})/w(\Delta_{t=1}). \tag{4}$$

The effective discount factor $\tilde{\beta}$ parameterizes the degree to which the relative thinker acts present-biased: It is the degree of present bias that an analyst with knowledge of $u(\cdot)$, I, k, and $\delta = 1$ would estimate to match the person's choice of consumption if he ignored relative thinking. Comparative statics on $\tilde{\beta}$ thus identify factors that shape the degree to which a relative thinker appears present-biased. Since $\Delta_{t=2}$ is increasing in k while $\Delta_{t=1}$ is constant in k, the person's effective discount factor is decreasing in k. Summarizing:

Proposition 4. *Defining $\tilde{\beta}$ as in Equation* (4)*, then $\tilde{\beta} \leq 1$ and $\partial \tilde{\beta}/\partial k < 0$.*

Proposition 4 says that the person acts as if he discounts the future by more when future consumption utility is more uncertain. The intuition is that the range of second-period consumption utility goes up with such uncertainty while the range of first-period consumption utility stays the same, leading a change in second-period consumption utility to loom smaller relative to a change in first-period consumption utility.[37]

[36] The maximum of the range is achieved through the lottery F associated with spending nothing in the first period and equals $E_F[u] + 1/2 \cdot S_F[u] = u(I) + 1/2 \cdot k$. The minimum of the range is achieved through the lottery G associated with spending everything in the first period and equals $E_G[u] - 1/2 \cdot S_G[u] = 0 - 1/2 \cdot k = -1/2 \cdot k$.

[37] With the simplifying assumption that the person cannot borrow, greater *income* uncertainty also robustly leads the relative thinker to discount the future by more. However, if greater future income uncertainty restricts the person's ability to borrow in the first period then it is ambiguous how it affects the degree to which the relative thinker discounts future consumption: Such uncertainty will increase the minimum of the range of second-period consumption utility, pushing the relative thinker to discount the future by less, and decrease the range in first-period consumption utility, pushing the relative thinker to discount the future by more. Which effect wins out depends on the curvature of the utility function. The latter effect is stronger with linear consumption utility, for example.

We now consider consumption over more than two periods: $T > 2$. Several issues need to be sorted out to make predictions in this setting. As an illustration, assume a person's instantaneous consumption utility in each period t is $\sqrt{c_t}$ and there is no uncertainty. Imagine he has \$10,000 to spend over 100 days and he is deciding how much money to take out of an ATM machine to spend today. Suppose for now that, while at the ATM, the person treats each period as a separate consumption dimension. With these assumptions, the range of consumption utility equals $\sqrt{10000} - \sqrt{0} = 100$ along each of the 100 dimensions and no present bias will be induced. But should we really expect separate days to be treated as separate dimensions? We think an alternative assumption is more realistic. In particular, suppose a person at the ATM treats "today" as one dimension and integrates all other periods together into a single "future" dimension. Then the range of consumption utility on the today dimension remains 100, but the range of consumption utility on the future dimension becomes larger: assuming that the maximum of the range comes from optimally spending \$10,000 over 99 days and the minimum comes from (optimally spending) \$0 over 99 days, then the range equals $99 \cdot \sqrt{10000/99} - \sqrt{0} = 300 \cdot \sqrt{11}$. As a result, the person effectively discounts the future by factor $\tilde{\beta} = w(300 \cdot \sqrt{11})/w(100) < 1$, and acts present-biased.

The above example illustrates the importance of the degree to which relative thinkers think about consumption at different periods as different dimensions. Kőszegi and Szeidl (2013) and the application of their model by Canidio (2014) assume that a person treats each period as a separate consumption dimension, and by examining natural cases where choices have a big impact on immediate utility but a dispersed impact on future utility, their model of overweighting big dimensions induces present bias. We think a more natural assumption is that the person thinks more precisely about how he spends money today than tomorrow. Kőszegi and Szeidl-style focusing induces "future-bias" if a person indeed integrates future consumption, whereas relative thinking induces present bias.[38] Rather than being a novel formulation, we view our assumption that people combine future spending into one aggregate "utility-of-spending dimension" as being implicit in other weighted-dimension models that treat "money" as a single dimension in static implementations.[39]

Even assuming as we do moving forward that a person treats the future as one dimension, the

[38] As such, one way to separate out the predictions might be to examine the impact of framing manipulations aimed at influencing the extent to which the person segregates future consumption. If asking a person to write down how much he plans to spend in each of the 100 days induces him to increase spending, that would support the Kőszegi and Szeidl model; we predict this will induce the person to spend less today. Because other factors may be at play, however, we hesitate to say that observing such a decrease would be a clean vindication of our model.

[39] Indeed, like us, Kőszegi and Szeidl (2013) and Bordalo, Gennaioli, and Shleifer (2013) do not segregate money or price over multiple dimensions when they consider static versions of their models. This matters for some predictions. For example, Koszegi and Szeidl say that people necessarily maximize consumption utility when choosing whether or not to buy a single-dimensional good at a given price since this constitutes a "balanced choice" when money is treated as a single dimension. This result would not continue to hold under the alternative assumption that money could be treated as spread out over multiple dimensions.

implications of relative thinking on spending depend on how much future utility a person imagines that his savings will bring him. We suppose the consumer treats the consumption value of savings as if he spends it optimally. Formally, we assume that, in period t a person values having I_{t+1} to spend from tomorrow on at $V_{t+1}(I_{t+1}) = \max_{c_{t+1},\ldots,c_T} \sum_{k=t+1}^{T} u(c_k)$, where the c_k satisfy $\sum_{k=t+1}^{T} c_k = I_{t+1}$. This is a form of naivete when there is more than one period left because our model predicts he will *never* spend in a way that maximizes total consumption utility. The naivete we assume implies that the person will norm future utility according to a range determined by \$0 future spending (as if he spent everything today) to a maximum of this optimal spending. We suspect that our qualitative results do not depend on naivete, and that the much harder to solve and harder to fathom alternative of sophisticated behavior based on the recursively constructed actual spending will yield very similar results.

As in the two period case, we assume each future period's consumption utility is uncertain. Utility in period τ is given by $u(c_\tau) \pm k_\tau$, $k_\tau \geq 0$, and the 50/50 gain k_τ, lose k_τ lottery is realized in period τ. From today's perspective, the range of consumption utility on the "today" dimension is then $u(I_t) - u(0) = u(I)$ and the range of consumption utility on the "future" dimension is $V_{t+1}(I_t) - V_{t+1}(0) + S(\sum_{\tau=t+1}^{T} \pm k_\tau) = (T-t)u\left(\frac{I_t}{T-t}\right) + S(\sum_{\tau=t+1}^{T} \pm k_\tau)$, where $S(\cdot)$ again refers to average self-distance. In period t the relative thinker then chooses the c_t that solves

$$\max_{c_t \in [0, I_t]} w(u(I_t)) \cdot u(c_t) + w\left(V_{t+1}(I_t) + S\left(\sum_{\tau=t+1}^{T} \pm k_\tau\right)\right) \cdot V_{t+1}(I_t - c_t).$$

The person in period t chooses consumption c_t as if he were a naive quasi-hyperbolic (β, δ) discounter with $\delta = 1$ and "effective β" equal to

$$\tilde{\beta}_t \equiv \frac{w\left(V_{t+1}(I_t) + S\left(\sum_{\tau=t+1}^{T} \pm k_\tau\right)\right)}{w(u(I_t))}, \tag{5}$$

so assuming an interior solution the person chooses the c_t that solves[40]

$$u'(c_t) = \tilde{\beta}_t V'_{t+1}(I_t - c_t) = \tilde{\beta}_t u'\left(\frac{I_t - c_t}{T-t}\right).$$

[40]Our analysis is unrealistic in assuming no "true" quasi-hyperbolic discounting (e.g., Laibson 1997, O'Donoghue and Rabin 1999). With $(\beta, \delta) = (\beta, 1)$, the range of consumption utility the person attaches to different temporal dimensions can differ even if he assesses separate days as separate dimensions, but the degree to which the person integrates consumption in future periods still matters for predictions. Returning to the above example, under the assumption that the person segregates all future consumption dimensions, then the range of consumption utility in future periods is $\beta \cdot 100 < 100$, so range-based relative thinking counteracts present-bias (but does not reverse it under *N2*). Conversely, under our assumption that the person integrates together all future consumption, the range of consumption utility in future periods is $\beta \cdot 300 \cdot \sqrt{11}$, so range-based relative thinking reinforces present-bias.

The effective $\tilde{\beta}_t$ parameterizes the degree to which the relative thinker acts present-biased in period t.

Proposition 5. *Define $\tilde{\beta}_t$ as in Equation* (5)*, and suppose $t \leq T-1$ and $I_t > 0$. Then:*

1. *$\tilde{\beta}_t \leq 1$ with equality if and only if $T-t=1$ and $k_T = 0$.*
2. *Fixing I_t, $\tilde{\beta}_t$ is strictly decreasing in $(T-t)$.*

Part 1 of Proposition 5 says that the relative thinker acts present biased in all periods except the second-to-last. If there is no uncertainty in the final period, the relative thinker maximizes consumption utility given remaining income by spreading that income evenly over the final two periods. If there is uncertainty in the final period, the relative thinker acts present biased in all periods. Part 2 says that, fixing remaining income, he acts more present biased the more periods remain. These results follow from the fact that the degree to which the range of consumption utility is larger in future periods than today is increasing in the number of remaining periods. This is for two reasons: a greater number of remaining periods increases (i) future uncertainty and (ii) the utility of being able to spend a fixed amount in the future.

These results identify some factors that influence the degree to which people appear present-biased. Under certain conditions, people act less present-biased if they are frequently allotted a small amount to consume over a few periods than if alloted a proportionally larger amount to consume over many periods, and will act more present biased towards the beginning than the end of a consumption window. To illustrate, Proposition 5 tells us that, abstracting from uncertainty, a person who is given $2y > 0$ to consume over two days will consume y each day, while a person who is given $4y$ to consume over four days will consume more than y on the first day.

While difficult to test empirically, these results connect to evidence of "first-of-the-month effects", where people appear to spend more on "instantaneous consumption" (e.g., fresh food or entertainment) following monthly receipt of income (Stephens 2003, 2006; Huffman and Barenstein 2005) or government benefits like food stamps (Shapiro 2005; Hastings and Washington 2010). Noting that the decrease in consumption over the course of the benefit month is calibrationally difficult to explain with models of exponential discounting, Shapiro (2005) and Huffman and Barenstein (2005) argue that the patterns are instead reasonably consistent with models of present-bias.[41] Our model is consistent with this evidence, but—under narrow bracketing assumptions that take T to be the end of a benefit month—makes the further prediction that households may appear less present-biased in their consumption decisions towards the end of the benefit month, and also if government benefits or paychecks were distributed more frequently. In our model, a more frequent

[41] Indeed, Shapiro (2005) makes the policy recommendation that food stamps should be distributed more frequently and in smaller amounts.

allocation would not only give people less of an opportunity to misbehave but could additionally limit their desire to misbehave. To the best of our current knowledge, these predictions have not been explored.[42]

We now turn to another aspect of intertemporal choice that may be heavily influenced by relative thinking: Situations in which a person can spend on multiple consumption goods each period. Our key assumption here is that the person continues to segregate consumption of goods that he considers buying today but to integrate the consumption of goods he will buy in the future. When the person is deciding whether to buy a candy bar for price $p at a store he does not precisely think about what else he could consume with $p tomorrow, but rather has a more abstract sense of the marginal utility of $p tomorrow. Our interpretation stresses when a good is purchased, not necessarily when it is consumed. For example, a person may purchase tickets to a concert that he goes to at a later date. As noted above, we believe that something like this distinction must underlie examples from the context-dependent literature where money is treated as a single dimension.

Appendix A.3 works out a simple two-period example in which a relative thinker overspends on goods he can buy in the first period. Since the relative thinker segregates the consumption of goods purchased today but integrates the consumption of goods purchased in the future, the range attached to the consumption of goods purchased today is smaller and the relative thinker attaches more weight to a given incremental change in consumption utility. A relative thinker will spend more on concert tickets that are released the same day he receives a cash windfall than later, even if he does not attend the concert until later.

6 Relationship to Focusing

As we saw above, the premise of salience and focusing models by Bordalo, Gennaioli, and Shleifer (2012, 2013) and Kőszegi and Szeidl (2013), namely that a wider range of outcomes draws people's attention to a dimension, can be in clear tension with the form of relative thinking we emphasize. The formulation by Kőszegi and Szeidl (2013) goes further by implying that wider ranges always increase people's sensitivity to fixed changes within the range. While we share the intuition behind the premise of the salience and focusing models, the Kőszegi and Szeidl (2013) formulation is directly in opposition to the evidence on range effects that motivates our model. As a result, we

[42]Recall that, as with quasi-hyperbolic discounting, a person may simultaneously act very present-biased in period t and not change his consumption by much between t and $t+1$. Consumption will likely drop off slowly towards the beginning of a budget window. Also, we only predict that the relative thinker will act less present-biased towards the end of a budget window with respect to trade offs among remaining days of that budget period. Because the person's marginal utility of income for instantaneous consumption will rise, he may favor income received at the end of a pay cycle quite heavily over future income. If a relative thinker were asked the question of the sort Shapiro (2005) analyzed, "What is the smallest amount of cash you would take today rather than getting $50 one month from today?", he would likely seem to be extremely impatient.

need to combine a different model of focusing with our model of relative thinking to study the net effect of expanding the range. This section sketches such a model.

As motivation, we believe some natural intuitions can guide speculation for when relative-thinking or focusing-effects may dominate. Our analysis above highlights examples where people's decisions are likely guided by a clear, low-dimensionality trade-off—between money and effort, money and quality, risk and return, etc. Per Figure 2 in Section 3, we believe that many of the sharp, direct predictions of range-based focusing models contradict evidence and intuition in two dimensions. But both our analysis and the evidence we provide may be misleading by "sampling" only from situations where people's attention is directed to the relevant dimensions. In situations where dimensions may be neglected, the bigger-range-increases-incremental-weight hypothesis might be the dominant force, perhaps by approximating the idea that people stochastically notice or pay attention to dimensions according to their range.[43] In this light, it is notable that most of the examples provided by Kőszegi and Szeidl address trade-offs across many dimensions—and that virtually none of our examples consider more than 3 dimensions. With more dimensions, it becomes more intuitive that people may concentrate their attention on dimensions with wider ranges to the point of paying more attention to incremental changes along those dimensions.

This suggests that focusing effects might naturally arise in situations with many dimensions. A crude formulation might hold that people pay attention to the two dimensions with the greatest ranges, and make choices according to range-based relative thinking, applied as if those are the only two dimensions.[44] Less crudely, we might suppose that other dimensions are partially attended to, but get decreased weight when their ranges are smaller. Consider the following formulation that channels this intuition, while maintaining the feature—both of our model and of Kőszegi and Szeidl (2013)—that people pay equal attention to dimensions when there are only two. First, 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 $(3, 2, 1)$

[43] In models of inattention such as Gabaix (2014) and Schwartzstein (2014), the range of potential realizations on some dimension of concern might naturally be a determinant of the likelihood that people pay attention.

[44] In contrast to Kőszegi and Szeidl (2013), the focusing part of our model is inspired by the idea that people have a limited "attentional budget", which is a central component of some models of rational inattention (e.g., Sims 2003).

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 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).

7 Discussion: Shortcomings of the Model

We believe that range-based relative thinking is one of the most acknowledged aspects of human perception and judgment, and corresponds to strong intuitions about the choices people make. Yet the connection between the basic psychology and the model we have developed, and, in turn, the economic implications, may not always be so sharp. We conclude the paper by discussing some of the shortcomings of the model, emphasizing missing elements and countervailing intuitions. We assess which conclusions from our model—as well as from models built around some of these countervailing intuitions—might be misleading, and suggest some potential for improvements.

A conspicuous omission from our model is reference dependence. This is most notable above and in Appendix C when we compare range-based relative thinking to diminishing sensitivity as embedded in Bordalo, Gennaioli, and Shleifer (2012, 2013). In order to incorporate reference dependence, extensions of our model must define the reference point—and must do so consistently across contexts—in order to study the interaction between relative thinking and reference-based effects like diminishing sensitivity and loss aversion. Bushong, Rabin, and Schwartzstein (2016, in progress) integrate relative thinking with reference-dependent preferences, following a variant of prospect theory along the lines of Kőszegi and Rabin (2007).[45] The model integrates the two in a particular way—norming both consumption utility and the sensations of gains and losses—that generates the prediction that exposure to bigger risks makes a person less risk averse. For example, if a homeowner is required to purchase some sort of insurance policy, then adding policies with higher deductibles and lower premia to an existing menu can only lead a person to choose higher

[45] The results in Bushong, Rabin and Schwartstein (2016, in progress) focus more on how relative thinking interacts with loss aversion than on how it interacts with diminishing sensitivity, but the model could also be used to study the latter question.

deductible policies. The analysis sheds light on evidence of certain context effects in risky choice, such as Post, van den Assem, Baltussen, and Thaler (2008), whereby people display less risk aversion in small to medium stakes decisions when they expect to have the option to take risks that involve bigger stakes.

Our model shares a set of limitations with other models of context-dependent preferences.[46] Like other models, ours takes as given the options a person considers at the time of choice. As the jacket-calculator example from earlier sections illustrates, our explanation hinges on people narrowly bracketing spending on a given item. While such narrow bracketing—stressed by authors such as Tversky and Kahneman (1981); Benartzi and Thaler (1995); Read, Loewenstein, and Rabin (1999); Barberis, Huang, and Thaler (2006); Barberis and Huang (2009); Rabin and Weizsacker (2009) and others—is consistent with the psychological evidence, there is little by way of general and systematic analysis of the extent, patterns, and implications of bracketing. Although we discussed in Section 5 some of the issues of bracketing and the entangled issues of how people integrate dimensions, the analysis there also illustrates the centrality of such assumptions to predictions, and highlights the need to develop more satisfactory models.

Perhaps a more fundamental limitation is one that pervades the entire choice-set-dependent literature: a near-silence on the question of what exactly lies in the "comparison set". Rarely are the "comparison sets" posited in the literature truly the choice sets that people face. For example, we can buy a car with or without a car radio ... but we can also buy another car. We think the intuition underlying most examples in papers on context-dependent preferences clearly relies on reasonable notions of what a person "might" do, but this remains to be formalized more carefully. We have tried to exclude any examples where our predictions are sensitive to options that might reasonably be added to the comparison set, though it is hard to evaluate our success in doing so—or quite what success would mean.

Treating the comparison set as exogeneous can also give misleading impressions on how context-dependent preferences will impact choices in market situations, where firms have an incentive to influence the comparison set through the products that they market. In fact, we think the results on "prophylactic decoys" sketched above and formalized in Appendix A.2 should give us pause on how results about decoys are interpreted in our model—and in other context-dependent models. In certain market contexts, these results indicate that, even though firms with inferior products might be able to use decoys to draw business away from a passive firm's superior product, the superior firm would prevail in an equilibrium where all firms can choose decoys. This suggests the impor-

[46] One further omission is noteworthy since it is specific to our model. Our model is inspired by the psychological research underlying range-frequency theory (Parducci 1965), but does not include a notion of "frequency"—that is, on how norming can depend not only on ranges but also the distribution within the ranges. Such frequency effects may matter for choice, but we do not understand them very well, and suspect that such effects are largely separable from the range-based effects we emphasize here.

tance of considering the market structure, as well as the marketing and production technologies, in thinking about how context-dependent preferences influence market outcomes.

A final limitation of the model concerns not so much an omission in the framework necessary to make predictions, but rather a limitation to the questions asked. The analysis emphasizes the behavioral implications of the model. The welfare interpretation that seems most consonant with our presentation—that relative thinking influences choice but not hedonic utility—seems realistic in many situations. However, in some situations it seems plausible that the way choices are hedonically experienced depends on how they were normed. For example, in situations involving risk, the difference between losing \$10 and \$5 may *feel* smaller when losing \$100 was possible. The model does not provide guidance on when relative thinking reflects a mistake or corresponds to true experienced well-being.

References

Azar, Ofer H, "Do People Think About Absolute or Relative Price Differences When Choosing Between Substitute Goods?," *Journal of Economic Psychology*, 2011, *32* (3), 450–457.

Barberis, Nicholas and Ming Huang, "Preferences with Frames: A New Utility Specification that Allows for the Framing of Risks," *Journal of Economic Dynamics and Control*, 2009, *33* (8), 1555–1576.

_ , _ , and Richard H Thaler, "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *The American Economic Review*, 2006, pp. 1069–1090.

Baron, Jonathan, "Confusion of Relative and Absolute Risk in Valuation," *Journal of Risk and Uncertainty*, 1997, *14* (3), 301–309.

Benartzi, Shlomo and Richard H Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *The Quarterly Journal of Economics*, 1995, *110* (1), 73–92.

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).

_ , _ , and _ , "Memory, Attention and Choice," 2015. Mimeo, Harvard University.

Bracha, Anat and Uri Gneezy, "Relative Pay and Labor Supply," 2012. Mimeo, Boston Fed.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein, "Relative Thinking, Reference Dependence, and Risk Attitudes," 2016.

Canidio, Andrea, "Focusing Effect and the Poverty Trap," 2014.

Cole, Shawn, Xavier Gine, and James Vickery, "How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment," 2013. Mimeo, Harvard.

Cunningham, Tom, "Comparisons and Choice," 2013. Mimeo, Stockholm University.

den Assem, Martijn J Van, Dennie Van Dolder, and Richard H Thaler, "Split or Steal? Cooperative Behavior When the Stakes are Large," *Management Science*, 2012, *58* (1), 2–20.

Desvousges, William, F Reed Johnson, Richard W Dunford, Sara P Hudson, and K Nicole Wilson, "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability," *Contingent Valuation: A Critical Assessment*, 1993, p. 91.

Doyle, John R, David J O'Connor, Gareth M Reynolds, and Paul A Bottomley, "The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases," *Psychology & Marketing*, 1999, *16* (3), 225–243.

Fehr, Ernst and Antonio Rangel, "Neuroeconomic Foundations of Economic Choice: Recent Advances," *The Journal of Economic Perspectives*, 2011, *25* (4), 3–30.

Frederick, Shane and Baruch Fischhoff, "Scope (In) Sensitivity in Elicited Valuations," *Risk Decision and Policy*, 1998, *3* (2), 109–123.

_, Leonard Lee, and Ernest Baskin, "The Limits of Attraction," *Journal of Marketing Research*, 2014, *51* (4), 487–507.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4).

Hastings, Justine and Ebonya Washington, "The First of the Month Effect: Consumer Behavior and Store Responses," *American Economic Journal: Economic Policy*, 2010, *2* (2), 142–62.

Heath, Timothy B and Subimal Chatterjee, "Asymmetric Decoy Effects on Lower-Quality Versus Higher-Quality Brands: Meta-Analytic and Experimental Evidence," *Journal of Consumer Research*, 1995, pp. 268–284.

Herne, Kaisa, "Decoy Alternatives in Policy Choices: Asymmetric Domination and Compromise Effects," *European Journal of Political Economy*, 1997, *13* (3), 575–589.

Highhouse, Scott, “Context-Dependent Selection: The Effects of Decoy and Phantom job Candidates,” *Organizational Behavior and Human Decision Processes*, 1996, *65* (1), 68–76.

Hsee, Christopher K, Jiao Zhang, Zoe Y Lu, and Fei Xu, “Unit Asking: A Method to Boost Donations and Beyond,” *Psychological Science*, 2013, *24* (9), 1801–1808.

Huber, Joel and Christopher Puto, “Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects,” *Journal of Consumer Research*, 1983, pp. 31–44.

_ , John W Payne, and Christopher Puto, “Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis,” *Journal of Consumer Research*, 1982, pp. 90–98.

Huffman, David and Matias Barenstein, “A Monthly Struggle for Self-Control? Hyperbolic Discounting, Mental Accounting, and the Fall in Consumption Between Paydays,” 2005.

Jahedi, Salar, “A Taste for Bargains,” 2011. Unpublished Manuscript, University of Arkansas.

Janiszewski, Chris and Donald R Lichtenstein, “A Range Theory Account of Price Perception,” *Journal of Consumer Research*, 1999, *25* (4), 353–368.

Kahneman, Daniel and Amos Tversky, “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica*, 1979, pp. 263–291.

_ and Jack L Knetsch, “Valuing Public Goods: The Purchase of Moral Satisfaction,” *Journal of Environmental Economics and Management*, 1992, *22* (1), 57–70.

Kamenica, Emir, “Contextual Inference in Markets: On the Informational Content of Product Lines,” *The American Economic Review*, 2008, *98* (5), 2127–2149.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry, “Agricultural Decisions after Relaxing Credit and Risk Constraints,” 2013. Mimeo, Yale.

Kivetz, Ran, Oded Netzer, and V Srinivasan, “Alternative Models for Capturing the Compromise Effect,” *Journal of Marketing Research*, 2004, *41* (3), 237–257.

Kontek, Krzysztof and Lewandowski Michal, “Range-Dependent Utility,” 2012. Mimeo, Boston Fed.

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

— **and Matthew Rabin**, "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

— **and** — , "Reference-Dependent Risk Attitudes," *The American Economic Review*, 2007, pp. 1047–1073.

— **and** — , "Reference-Dependent Consumption Plans," *The American Economic Review*, 2009, *99* (3), 909–936.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, pp. 443–477.

Louie, Kenway, Lauren E Grattan, and Paul W Glimcher, "Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex," *The Journal of Neuroscience*, 2011, *31* (29), 10627–10639.

Mellers, Barbara A and Alan DJ Cooke, "Trade-offs Depend on Attribute Range.," *Journal of Experimental Psychology: Human Perception and Performance*, 1994, *20* (5), 1055.

Mobarak, Ahmed and Mark Rosenzweig, "Selling formal insurance to the informally insured," 2012. Yale University Economic Growth Center Discussion Paper.

O'Donoghue, Ted and Matthew Rabin, "Doing it Now or Later," *American Economic Review*, 1999, pp. 103–124.

Padoa-Schioppa, Camillo, "Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex," *The Journal of Neuroscience*, 2009, *29* (44), 14004–14014.

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

— , "Contextual Effects: A Range-Frequency Analysis," *Handbook of Perception*, 1974, *2*, 127–141.

Pettibone, Jonathan C and Douglas H Wedell, "Testing Alternative Explanations of Phantom Decoy Effects," *Journal of Behavioral Decision Making*, 2007, *20* (3), 323–341.

Post, Thierry, Martijn J Van den Assem, Guido Baltussen, and Richard H Thaler, "Deal or no Deal? Decision Making Under Risk in a Large-Payoff Game Show," *The American Economic Review*, 2008, pp. 38–71.

Pratkanis, Anthony R and Peter H Farquhar, "A Brief History of Research on Phantom Alternatives: Evidence for Seven Empirical Generalizations About Phantoms," *Basic and Applied Social Psychology*, 1992, *13* (1), 103–122.

Prelec, Drazen and George Loewenstein, "Decision Making over Time and under Uncertainty: A Common Approach," *Management Science*, 1991, *2* (7), 770–786.

_ , Birger Wernerfelt, and Florian Zettelmeyer, "The role of Inference in Context Effects: Inferring What you Want From What is Available," *Journal of Consumer Research*, 1997, *24* (1), 118–126.

Rabin, Matthew and Georg Weizsacker, "Narrow Bracketing and Dominated Choices," *The American Economic Review*, 2009, pp. 1508–1543.

Rangel, Antonio and John A Clithero, "Value Normalization in Decision Making: Theory and Evidence," *Current Opinion in Neurobiology*, 2012.

Read, Daniel, George Loewenstein, and Matthew Rabin, "Choice Bracketing," *Journal of Risk and Uncertainty*, 1999, *19* (1-3), 171–97.

Rubinstein, Ariel, "Similarity and Decision-Making Under Risk (Is There a Utility Theory Resolution to the Allais Paradox?)," *Journal of Economic Theory*, 1988, *46* (1), 145–153.

Savage, Leonard, *The Foundations of Statistics*, New York: John Wiley and Sons, 1954.

Schwartzstein, Joshua, "Selective Attention and Learning," *Journal of the European Economic Association*, 2014, *12* (6), 1423–1452.

Shapiro, Jesse M, "Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 2005, *89* (2), 303–325.

Simonsohn, Uri, "New Yorkers Commute More Everywhere: Contrast Effects in the Field," *Review of Economics and Statistics*, 2006, *88* (1), 1–9.

_ and George Loewenstein, "Mistake# 37: The Effect of Previously Encountered Prices on Current Housing Demand," *The Economic Journal*, 2006, *116* (508), 175–199.

Simonson, Itamar, "Choice Based on Reasons: The Case of Attraction and Compromise Effects," *Journal of Consumer Research*, 1989, pp. 158–174.

_ and Amos Tversky, "Choice in Context: Tradeoff Contrast and Extremeness Aversion.," *Journal of Marketing Research*, 1992.

Sims, Christopher A, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Soltani, Alireza, Benedetto De Martino, and Colin Camerer, "A Range-Normalization Model of Context-Dependent Choice: A New Model and Evidence," *PLoS Computational Biology*, 2012, *8* (7), e1002607.

Stephens, Melvin, """ 3rd of tha Month": Do Social Security Recipients Smooth Consumption Between Checks?," *American Economic Review*, 2003, pp. 406–422.

—, "Paycheque Receipt and the Timing of Consumption," *The Economic Journal*, 2006, *116* (513), 680–701.

Strotz, Robert Henry, "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 1955, pp. 165–180.

Thaler, Richard, "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior & Organization*, 1980, *1* (1), 39–60.

—, "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.

Tsetsos, Konstantinos, Marius Usher, and Nick Chater, "Preference Reversal in Multiattribute Choice.," *Psychological Review*, 2010, *117* (4), 1275.

Tversky, Amos, "Intransitivity of Preferences.," *Psychological review*, 1969, *76* (1), 31.

— **and Daniel Kahneman**, "The Framing of Decisions and the Psychology of Choice," *Science*, 1981, *211*, 453–458.

— **and Itamar Simonson**, "Context-Dependent Preferences," *Management Science*, 1993, *39* (10), 1179–1189.

Volkmann, John, "Scales of Judgment and Their Implications for Social Psychology.," 1951.

Wedell, Douglas H, "Distinguishing Among Models of Contextually Induced Preference Reversals.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1991, *17* (4), 767.

Wernerfelt, Birger, "A Rational Reconstruction of the Compromise Effect: Using Market Data to Infer Utilities," *Journal of Consumer Research*, 1995, pp. 627–633.

Woodford, Michael, "Inattentive Valuation and Reference-dependent Choice," 2012.

Yang, Sybil and Michael Lynn, "More Evidence Challenging the Robustness and Usefulness of the Attraction Effect," *Journal of Marketing Research*, 2014, *51* (4), 508–513.

Yitzhaki, Shlomo, "Stochastic Dominance, Mean Variance, and Gini's Mean Difference," *The American Economic Review*, 1982, *72* (1), 178–185.

A Further Definitions and Results

A.1 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 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

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

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

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

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 6. *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 6 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.[47] Thaler gives the

[47] Note that, in contrast to diminishing sensitivity of the prospect theory value function, Proposition 6 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 6 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.

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 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 6, the evidence on losses is ambiguous.

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

Proposition 7. *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.2 Further Results on the Limits of Choice-Set Effects and Prophylactic Decoys

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

Corollary 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 8. *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 8 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 2. *Assume the conditions of Proposition 8 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.[48] 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 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.[49]

A.3 Further Results on Making Tradeoffs Across Time

Simplify the intertemporal choice model to two periods and abstract from uncertainty. Suppose there are n goods, $g = 1, \ldots, n$, and the person can spend on goods $1, \ldots, j < n-1$ in the first period and on $j+1, \ldots, n$ in the second. For example the person may find herself in a store that only sells goods $1, \ldots, j$ in the first period and in a store that only sells the remaining goods in the second.

As motivated in Section 5.2, suppose the person segregates consumption of goods that he considers buying today but integrates the consumption of goods he will buy in the future. The range of consumption utility the person attaches to a good purchased today is then $u_g(I/p_g) - u(0) = u_g(I/p_g)$, while the range of consumption utility the person attaches to a good purchased tomorrow is $V(p,I,j) - V(p,0,j) = V(p,I,j)$. Here I denotes the person's income, p_g the price of good g, and $V(p,I,j) \equiv \max_{c_{j+1},\ldots,c_n} \sum_{g=j+1}^{n} u_g(c_g)$ subject to $I = \sum_{g=j+1}^{n} p_g c_g$ represents the indirect utility of money tomorrow from today's perspective.

[48] 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.

[49] 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.

The person then chooses consumption in the first period to solve

$$\max_{c_1,\dots,c_j} \sum_{g=1}^{j} w(u_g(I/p_g)) \cdot u_g(c_g) + w(V(p,I,j)) \cdot V\left(p, I - \sum_{g=1}^{j} p_g c_g, j\right),$$

subject to each $c_g \geq 0$ and $\sum_{g=1}^{j} p_g c_g \leq I$.

To simplify the analysis, suppose the goods are symmetric, so the price and utility functions of all goods are the same, where we normalize the common price to equal 1 and assume the common utility function satisfies $u(0) = 0, u'(\cdot) > 0, u''(\cdot) < 0$. In this case, absent relative thinking the person will spread consumption equally across all goods: $c_g = I/n$ for all g. What happens with relative thinking?

In the first period, the relative thinker attaches weight $w_1 \equiv w(u(I))$ to all goods purchased in the first period and weight $w_2 \equiv w((n-j) \cdot u(I/(n-j)))$ to all goods purchased in the second, where $w_2 < w_1$ under the assumption that $j < n-1$. The relative thinker thus looks impatient: he spends more than I/n on all goods purchased in the first period and less than I/n on all goods purchased in the second.

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 and 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* 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 implies that

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

because $w(\Delta) \cdot \Delta$ is strictly increasing in Δ. Dividing by p and letting $p \to 0$ yields

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

which gives $u_k'(x)$ (using the normalization that $u_1'(0) = 1$) and the entire utility function $u_k(\cdot)$ (using the normalization that $u_k(0) = 0$). 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 (Volkmann 1951, 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:[50]

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

[50] 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).

where the level of w_i depends only on the salience rank of attribute i for option c in comparison set C.[51] 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 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[52]

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

[51] 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.

[52] 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.

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 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.[53]

[53] As Bordalo, Gennaioli and Shleifer (2013) note, their model accommodates the attraction effect when peo-

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,

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 an (often implicit) 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.[54] 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

ple 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 $\{(70,-20),(80,-30),(80,-70)\}$.

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

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. In cases like this one where there is not a natural reference point, implicit-reference-point theories like Bordalo, Gennaioli, and Shleifer (2013) and Cunningham (2013) can have more degrees of freedom in explaining observed patterns of behavior.

D Proofs

Proof of Proposition 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

$$U^N(c|C) - U^N(c'|C) = \sum_j w(\Delta_j(C)) \left[u_j(c_j) - u_j(c'_j)\right] = \sum_j w\left(\frac{\delta_j(c,c')}{d_j(c,c'|C)}\right)\delta_j(c,c') \geq 0, \quad (6)$$

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 $\tilde{C}$.

■

Proof of Proposition 2. 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*.

■

Proof of Corollary 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 2.
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 2.

■

Lemma 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 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 3. 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

$$U^N((F_1,F_2)|\mathscr{F})-U^N((F_1-G_1,F_2+G_2)|\mathscr{F})=w(\Delta_1(\mathscr{F}))\cdot E[G_1]-w(\Delta_2(\mathscr{F}))\cdot E[G_2]\geq 0, \tag{7}$$

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 (7) 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 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 4. In text. ■

Proof of Proposition 5. From Equation (5), we have

$$\tilde{\beta}_t = \frac{w\left((T-t)u\left(\frac{I_t}{T-t}\right) + S\left(\sum_{\tau=t+1}^{T} \pm k_\tau\right)\right)}{w(u(I_t))}.$$

We establish points 1-2 in turn.

1. By strict concavity of $u(\cdot)$ and non-negativity of $S(\cdot)$, $(T-t)u\left(\frac{I_t}{T-t}\right) + S\left(\sum_{\tau=t+1}^{T} \pm k_\tau\right) \geq u(I_t)$ with equality if and only if $T-t=1$ and $k_T=0$. The result that $\tilde{\beta}_t \leq 1$ with equality if and only if $T-t=1$ and $k_T=0$ then follows from the fact that $w(\cdot)$ is strictly decreasing (*N1*).

2. By strict concavity of $u(\cdot)$ and Lemma 1 in Koszegi and Rabin (2007), $(T-t)u\left(\frac{I_t}{T-t}\right) +$

$S\left(\sum_{\tau=t+1}^{T} \pm k_\tau\right)$ is strictly increasing in $(T-t)$, which implies that $\tilde{\beta}_t$ is strictly decreasing in $(T-t)$ by *N1*.

■

Proof of Proposition 6. 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

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

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

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

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 7. From Condition (1) of Proposition 2 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 6. ■

Proof of Proposition 8. 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 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 8 (letting $\varepsilon = \lambda'$), with $\lambda' = \lambda - \eta$ for $\eta > 0$ small.

By Proposition 8, 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 $c^j \in C_{\lambda'}$. ■