

A Model of Focusing in Economic Choice*

Botond Kőszegi
UC-Berkeley

Ádám Szeidl
UC-Berkeley, NBER, CEPR

April 29, 2011

Abstract

We present a theory of individual choice in which the decisionmaker focuses more on, and hence weights more heavily, attributes in which the options in her consideration set are more different. Consistent with evidence on salience in monetary choices, our model predicts a bias toward options whose advantages are concentrated in fewer attributes. In intertemporal choice, because the relative concentrations of an option's costs and benefits can be different from the perspective of a single period and the perspective of the entire choice problem, our model often predicts a form of time inconsistency. The decisionmaker exhibits present bias when the costs of current misbehavior are distributed over many future dates (such as in harmful consumption), but is "future-biased" when reaching a goal with concentrated benefits requires many periods of effort (such as in career advancement). In product design, a profit-maximizing firm chooses products with one core attribute and many price components. A strong firm wants to be especially strong on its competitor's weak attribute, while a weak firm wants to be relatively strong on its competitor's strong attribute.

*Takeshi Murooka provided excellent research assistance. We thank David Ahn, Stefano DellaVigna, Erik Eyster, Michael Grubb, Emir Kamenica, Joan de Martí Beltran, Matthew Rabin, and Richard Thaler for insightful discussions. Kőszegi thanks the Center for Equitable Growth and the National Science Foundation (Award #0648659) and Szeidl thanks the Alfred P. Sloan Foundation for financial support.

1 Introduction

When making a choice, individuals often do not focus equally on all attributes of their available options. An employee considering a financially attractive job offer may focus more on the large increase in yearly income than on the extra hours she will need to work every day (Kahneman, Krueger, Schkade, Schwarz and Stone 2006). Similarly, a person contemplating whether to live in California or the Midwest may focus more on the differences in climate and less on the many determinants of life satisfaction in which the two regions are similar (Schkade and Kahneman 1998). As psychologists have emphasized, this kind of focusing can generate supoptimal decisions if the person overweights the attributes she focuses on.

Motivated by examples and evidence such as these, we develop a model of focusing based on the idea that a person focuses more on attributes in which her options differ more. While this is only one of several possible determinants of focus, it seems relevant in many settings: our model explains under one framework a number of facts about salience, reconciles seemingly conflicting evidence and intuitions about intertemporal choice and product design, and also makes a number of new predictions in these domains. Because our model makes a prediction in any choice problem in which a classical model would, it also provides a simple way to incorporate this type of focusing into other economic applications.

Section 2 presents our framework. We model choices from a finite consideration set C , which we allow to be different from the agent’s entire choice set because there may be some available options (such as extremely unattractive options) that do not affect focus. Each option $c \in C$ provides utility on K different “attributes,” with the *consumption utility* from a choice—identical to welfare in our model—being equal to $U(c) = \sum_{k=1}^K u_k(c)$.¹ But instead of consumption utility, the decisionmaker acts to maximize *focus-weighted utility* $\sum_{k=1}^K g_k u_k(c)$, where the g_k weight captures the extent to which she focuses on attribute k . A key challenge in incorporating focusing into economics is to define focus—here, the weights g_k —without appealing to “outside” information that is difficult to observe or not relevant in most economic situations. We address this challenge by assuming that

¹ In this paper, we consider only riskless choices. In Kőszegi and Szeidl (2011, hopefully), we study an extension that allows for uncertainty by thinking of ordered lottery outcomes as attributes.

the decisionmaker focuses disproportionately on attributes in which her options generate a greater range of consumption utility: $g_k = g(\Delta_k(C))$, where $\Delta_k(C) = \max_{c \in C} u_k(c) - \min_{c \in C} u_k(c)$ and $g(\cdot)$ is an increasing function. While our basic questions, applications, and results are different, our basic theory parallels the models for study choice under uncertainty developed by Bordalo, Gennaioli and Shleifer (2010), Bordalo (2011), and Gabaix (2011), in which salience depends on the variation in potential outcomes.² We extend our basic setup to choice over time by assuming that a person thinks of her choice set in a period as the set of lifetime consumption profiles made possible by current choices, given her beliefs about future behavior.

In Section 3, we state and discuss the central prediction of our model: that the decisionmaker exhibits a “bias towards concentration”. Because she focuses too much on a few large advantages relative to many small disadvantages, a person is too prone to choose options whose advantages are concentrated in fewer attributes. This prediction—equivalent to our main assumption that $g(\cdot)$ is increasing—helps organize several findings about salience in monetary choices. First, if transfers on different dates are separate attributes, our model predicts that a person feels more strongly about one big payment than about a sequence of smaller payments. This can help explain why most taxpayers choose one lump-sum tax refund over lower monthly withholdings (Jones 2010), why retirees seem to take too much of their retirement wealth in a lump sum rather than as an annuity (Brown 2009), and why financing consumer for products—which replaces one big payment with a sequence of smaller ones—is so popular. Second, if different components of a price are separate attributes, our model predicts that consumers will perceive a product to be less expensive if its price is split into multiple components. This can help explain my demand for a taxable item in the supermarket is lower if its tax-inclusive price is posted on the shelf (Chetty, Looney and Kroft 2009), and why retailers (e.g. airlines and mortgage brokers) often artificially split prices into multiple components.

Importantly, the bias towards concentration is not operational in “balanced choices”—decisions in which the advantages and disadvantages of one option relative to another are equally concentrated. As a result, in balanced choices a person maximizes consumption utility. This observation

² We discuss direct evidence for our model and these related approaches in more detail in Section 2 below.

helps identify situations in which choices reflect welfare, and allows us to elicit consumption utility and other model ingredients from choice behavior. In this sense, our model satisfies the spirit of the revealed-preference criterion for economic theories.

In Section 4, we explore the implications of our model for intertemporal choice. Assuming that utility outcomes on different dates correspond to different attributes, the central principle emerging from our model is that the agent’s weighting of an option’s costs and benefits depends not on their temporal placement, but on their *concentration*. Furthermore, because the concentration of costs and benefits can be different from the ex-post perspective of a single choice and the ex-ante perspective that aggregates multiple decisions, the decisionmaker often exhibits a kind of time inconsistency. Our model predicts both how the choice problem shapes the strength and direction of time inconsistency, and whether the ex-ante or ex-post preference more accurately reflects welfare.

As an example of one class of decisions, consider the daily choice of whether to exercise, where doing so results in pain today and small benefits for many periods in the future. Because on each day a person focuses too little on the many small gains relative to the one big cost, she exhibits overly present-oriented behavior. But when considering the entire sequence of decisions ex ante, the benefits accumulate into larger and hence more salient gains, so that the person would commit to more future-oriented behavior. This scenario fits many types of decisions, including lifestyle choices such as diet and consumption-savings decisions, that have been invoked as prime examples of present bias in the literature. In these settings, our model both explains present bias and predicts that—it being a more balanced choice—the ex-ante decision more accurately reflects welfare.

As an example of a different class of decisions, consider choices, such as writing a book, where a sequence of sacrifices leads to a single large goal. Because ex ante the decisionmaker focuses too much on the large goal, she is too prone to commit to writing the book. But on any given day ex post, the person weighs effort that day against just a marginally better book, making her less willing to work. This scenario fits career and other “achievement”-type decisions, for which some researchers believe that individuals commit themselves to overly future-oriented paths (Scitovsky 1976, Kahneman et al. 2006, Hamermesh and Slemrod 2008). In this case, our theory says that—it being the balanced choice—the *ex-post* decision more accurately reflects welfare. Nevertheless, since

the diffuse incremental costs associated with each *additional* achievement look small relative to the single benefit, we also predict that the decisionmaker often chooses to add ever more consumption-utility-decreasing challenges on top of her current commitments. As these examples show, our model simultaneously explains common examples of present-oriented and future-oriented behavior, and also provides a simple welfare guide based on the relative balance of costs and benefits in the ex-ante and the ex-post decisions. Although we believe that time-inconsistent preferences often arise because of hyperbolic discounting and related mechanisms, we also believe that time inconsistency based on the shifting breadth of thinking captures an important aspect of intertemporal choice in practice.

In Section 5, we explore how firms design products when consumers have consideration-set-dependent focus, thereby contributing to a recent literature (reviewed in Spiegler 2011) that analyzes traditionally marketing questions using economic methods. Our model predicts that a monopolist evenly divides the price among available payment components, while it concentrates product value on a single attribute. The former implication is consistent with the retailer practice of splitting prices discussed above, and the latter seems central in marketing analyses of successful brand positioning and consumer “value propositions” (Aaker 1991, Anderson, Narus and van Rossum 2006, and others).

We also analyze firms’ incentives in choosing which value attribute to concentrate on. As exemplified by the association of Volvo with safety and of BMW with performance, firms often try to position their product by finding a single attribute in which they can dominate the competition. But as illustrated by the flood of touch-screen phones following the introduction of the iPhone, another common positioning strategy is to copy the key attribute of a competitor. Our model predicts that these two strategies are optimal for different kinds of firms. For a strong firm, which produces higher qualities and hence determines consumers’ focus, concentrating on the competitor’s weak attribute is optimal, as this minimizes focus on its competitor’s stronger attribute. For a weak firm, which does not affect consumer focus, concentrating on the competitor’s strong attribute is optimal, as this is the attribute that consumers focus more on.

While for the bulk of the paper we take the consideration set C as exogenous, in Section 6

we propose a theory of endogenous consideration-set determination that is consistent with our central assumption that differences attract attention. We posit that the decisionmaker considers the subset of her choice set that maximizes a linear combination of average consumption utility and average differences between options. We illustrate the use of our theory through an example. If a consumer’s choice set is whether or not to buy a particular item, unless the item is very expensive or very cheap her consideration set will be equal to her choice set. But if her choice set includes two possible attractive but different purchases, not buying might not make it into her consideration set. With her decision effectively shifting from *whether* to buy to *what* to buy, differences in money outlays become smaller, and hence the weight she puts on the money dimension diminishes. This can lead her to more easily buy a more expensive product than when there is only one product for purchase.

Especially with the endogenous theory of consideration sets, our model is highly portable: applying it to a new (deterministic) economic setting only requires—besides assumptions a classical model would also make—the specification of relevant attributes for that setting. This specification is fairly straightforward in many settings, and we also develop ways to use our predictions to identify a choice’s relevant attributes from behavior. We discuss other potential applications of and further issues with our model in Section 7.

2 A Theory of Focusing

We now formulate our model of focusing, first specifying a static version in Section 2.1, and then extending it to dynamic choice in Section 2.2. In this paper, we model only riskless choice; in Kőszegi and Szeidl (2011, hopefully), we extend our model to risky choice building on the approaches of Bordalo et al. (2010) and Bordalo (2011).

2.1 Focus-Weighted Utility

We model choices from a finite set $C \subset \mathbb{R}^K$ of K -dimensional consumption vectors, where each of the dimensions represents an “attribute.” Because it seems reasonable to assume that some available options, especially extremely and manifestly unfavorable ones, may not affect the decisionmaker’s

focus, we allow C to be different from her entire choice set, and think of C as a “consideration set” consisting of her reasonable options. While for the bulk of the paper we take C as exogenous, in Section 6 we propose a way to define it endogenously. Our theory of choice from consideration sets can also be combined with any other model of consideration-set determination.

The decisionmaker’s *consumption utility* from option $c = (c_1, \dots, c_K) \in C$ is $U(c) = \sum_{k=1}^K u_k(c_k)$, and we will often represent an option by its vector $(u_1(c_1), \dots, u_K(c_K))$ of consumption utilities in the K attributes rather than by its vector of consumption levels. When making choices, however, the decisionmaker maximizes *focus-weighted utility*

$$\tilde{U}(c, C) = \sum_{k=1}^K g_k \cdot u_k(c_k), \tag{1}$$

where g_k is the focus weight on attribute k .³ While focus-weighted utility determines choice, we assume that the decisionmaker’s welfare is given by her consumption utility. This assumption reflects our view (consistent with the psychology literature below) that focusing distorts the individual’s perception of preferences and perhaps her well-being near the moment of choice, but does not alter the much more important experienced utility emanating from the choice. In the words of Schkade and Kahneman (1998), “nothing in life is quite as important as you think it is while you are thinking about it.”

The major challenge for any applicable economic model of attention or focusing is to define focus—in our setting, the weights g_k —without appealing too much to determinants that are either difficult to observe or not relevant in most economic situations. For instance, both because data on this would be difficult to obtain and because it would be undefined for many economic questions, we do not want to assume that the decisionmaker focuses on attributes that are highlighted in red on the product’s packaging. Using only model ingredients introduced so far, our central assumption is that the decisionmaker’s focus on an attribute depends on how much her options’ consumption

³ Literally interpreted, Equation (1) says that focusing more on one attribute does not reduce focus on other attributes. But choice depends on the *relative* focus across attributes, and increasing the weight on one attribute does reduce the relative weight on other attributes. To make the role of relative weights more transparent, we could normalize each attribute’s weight by $\sum_{j=1}^K g_j$, which would ensure that the weights sum to one. This adjustment would merely multiply $\tilde{U}(c, C)$ by a constant, and hence would not affect the choice or welfare implications of the model.

utilities differ in that attribute:⁴

$$g_k = g(\Delta_k(C)), \text{ where } \Delta_k(C) = \max_{c \in C} u_k(c) - \min_{c \in C} u_k(c). \quad (2)$$

Within this class of decision utilities, much of our paper will make the following further basic assumption:

Assumption 1. The function $g(\Delta)$ is strictly increasing in Δ .

Assumption 1 says that the decisionmaker focuses more on attributes in which her options generate a greater range of consumption utility. Some evidence directly supports this assumption.⁵ For example, Dunn, Wilson and Gilbert (2003) examine freshmen’s predicted and upperclassmen’s actual level of happiness with their randomly assigned dorms at a major university. Consistent with our hypothesis, predicted happiness depends greatly on features (e.g. location) that vary a lot between dorms, and not on features (e.g. social life) that vary little between dorms—whereas actual happiness does not show the same patterns. Similarly, Schkade and Kahneman (1998) find that both Midwesterners and Southern Californians incorrectly predict Californians to be more satisfied with life because they focus on the main differences (climate and cultural opportunities) between the two locations and underweight important other determinants of life satisfaction. But rather than the direct evidence, we view the evidence below on the “bias toward concentration”—a prediction of our model that is equivalent to an increasing $g(\cdot)$ —as the most compelling support for our framework.

Several comments regarding our basic theory of focusing are in order. First, given the psychological evidence and economic examples motivating our model, the most natural interpretation of our formulation is that the decisionmaker is at some level aware of the attributes of her options, but does not focus equally on them. We refer to this phenomenon as “focusing” partly to distinguish our model from theories of attention in economics, where the issue is more about what information

⁴ Our formulation uses a particular tractable functional form to capture dispersion in consumption utility. We are not aware of evidence that would allow us to tie down this functional form precisely. The logic of our results in this paper, however, does not seem to depend on the specific formulation.

⁵ The evidence we discuss pertains to predicted happiness rather than choice—but presumably individuals would make choices in the above situations corresponding to their predicted happiness.

a person becomes aware of or understands. Nevertheless, when it seems semantically appropriate and does not cause confusion, we also refer to the phenomenon as attention.

Second, to simplify the definition of the range of consumption utility in an attribute (which in turn determines focus), our formulation assumes additive separability. None of the intuitions in our paper seem to depend on this simplification. Nevertheless, we offer a simple way to extend our definition to non-separable utility. Suppose that consumption utility is given by the potentially non-separable function $U(c_1, \dots, c_K)$. We posit that there is a “yardstick” option $c^0 \in C$, which can be taken as exogenous or defined as the (generically unique) consumption-utility-maximizing option. We let $\Delta_k(C) = \max_{c \in C} U(c_k, c_{-k}^0) - \min_{c \in C} U(c_k, c_{-k}^0)$, and define the focus-weighted utility of c with respect to c^0 as $wU(c) + (1 - w) \sum_{k=1}^K g(\Delta_k(C)) [U(c_k, c_{-k}^0) - U(c^0)]$. The first term allows complementarities in consumption utility to influence the agent’s behavior, and the second term captures our hypothesis that greater differences on an attribute lead the agent to overweight that attribute. Our basic model obtains as a special case when U is additively separable and $w = 0$.

Third, although the basic questions, applications, and results are different, our formulation of the decisionmaker’s behavior is related to recent theories of salience and attention in choice under uncertainty. In their model of choice between two lotteries, Bordalo et al. (2010) assume that the salience of a state of the world increases in the difference of the two lotteries’ payoffs in that state. This framework yields a salience-based foundation for the probability-weighting function in prospect theory, and has a number of new predictions for risky choice. Bordalo (2011) extends the framework to choice among multiple lotteries, and shows that the more general model parsimoniously explains context effects in choice behavior such as the compromise effect and asymmetric dominance. Gabaix (2011) assumes that the agent underweights or ignores factors in her decisions which, given the uncertainty in that factor in the environment, do not affect her utility very much.

Fourth, although our model’s ingredients are not based directly on choice behavior, we show in Appendix B that if we know the relevant attributes, all ingredients of our model can be identified from choices. This means that our theory’s full set of testable predictions as well as its welfare implications can be identified from behavior. In this sense, our model satisfies the spirit of the

revealed-preference criterion for economic theories. We also illustrate in the Appendix that if know the relevant attributes of some options, under some conditions we can identify the relevant attributes of other options.

2.2 Focusing in Intertemporal Choices

To explore the implications of focusing in intertemporal decisions, we now formulate a dynamic model. Our starting assumption is that outcomes at different dates correspond to different attributes. Even given this basic assumption, application of our framework requires us to specify how a person conceptualizes her dynamic choice problem. Our key assumption is that the agent represents her consideration set in a period as the set of lifetime consumption profiles associated with her current options, given her beliefs regarding her future behavior. In a consumption-savings decision, for instance, the individual has beliefs regarding how a dollar consumed today affects consumption on each future date. Through these beliefs, today’s consumption possibilities generate a set of lifetime consumption profiles, and this set determines the individual’s focus. Our formulation reflects the idea, broadly consistent with evidence on narrow bracketing (e.g. Tversky and Kahneman (1981) and Rabin and Weizsäcker (2009)), that—while they may have an idea about the future consequences of their actions—individuals make decisions one-by-one rather than by integrating them into a grand lifetime decision problem.

Formally, there are T periods, $t = 1, \dots, T$. In period t , the agent makes a choice c_t from the deterministic finite consideration set $C_t(c_{t-1}, \dots, c_1)$. Continuing with the consumption-savings example, c_t can represent the bundle of goods to be consumed today, and $C_t(c_{t-1}, \dots, c_1)$ the budget set given the path of past consumption and possible credit constraints. We abstract away from any time inconsistency or time discounting in intrinsic preferences, and assume that the decisionmaker’s consumption utility in period t is $\sum_{s=t}^T u_s(c_s, c_{s-1}, \dots, c_1)$ where u_s is the possibly history-dependent instantaneous utility function in period s . Besides requiring that utilities realized on different dates are evaluated as separate attributes, we also allow for multiple attributes at each date t , and assume that the instantaneous utility function is additively separable over these attributes.

The decisionmaker’s beliefs about the consequences of her current choice are captured by the functions $\{c_\tau(c_t)\}_{\tau=t+1,\dots,T}$, which specify future choices as a function of c_t . For any history c_1, \dots, c_{t-1} , the decisionmaker’s beliefs induce a set of lifetime consumption profiles from the current consideration set:

$$\bar{C}_t = \{(c_t, c_{t+1}(c_t), \dots, c_T(c_t)) \mid c_t \in C_t(c_{t-1}, \dots, c_1)\}.$$

We assume that \bar{C}_t is the consideration set that determines the decisionmaker’s focus in period t , so that she applies the model of Section 2.1 to \bar{C}_t .

Our framework can be combined with any theory of the agent’s beliefs regarding future behavior. Following standard economic methodology, in this paper we assume that the agent has rational (correct) beliefs about her future behavior. This implies that we can derive behavior in any decision problem using backward induction. In Appendix C, we formulate a “naive” theory, in which the decisionmaker believes that she will act in the future as she would commit to initially. The insights in this paper hold for both theories of how the agent forms beliefs about the future.

A key question we explore in intertemporal choice is whether the agent’s behavior is identical to what she would commit to *ex ante*. The *ex-ante* or commitment problem is a choice problem in which the agent makes all decisions at time 1, choosing from the set C_{ante}^* of all lifetime consumption profiles. Applying the dynamic model of focusing above, in this case (and in contrast to the no-commitment problem) the individual’s focus in period 1 is determined by the range of consumption utilities generated by all possible consumption paths.

3 Bias Toward Concentration

This section identifies the key implication of our model: that the decisionmaker is biased toward options whose advantages are more concentrated than its disadvantages. As a simple example, consider an employee who currently has \$3,000 per month in take-home pay, and receives a lump-sum tax refund of \$2,400 when filing her taxes the next year. She could ask her employer to reduce her withholding by \$200, which would result in \$3,200 per month in take-home pay and a \$0 tax refund. Does she do so? Assuming no discounting, a zero interest rate, linear consumption utility, and

positing that payments at different dates correspond to different attributes, the agent represents the decision as choosing between the streams $(3000, \dots, 3000, 2400)$ and $(3200, \dots, 3200, 0)$. Because the sum of payments in the two options is the same, a rational agent would be indifferent. In our model, however, the individual assigns a smaller focusing weight $g(200)$ to the first twelve monthly payments in which the difference between her options is only 200, and a larger weight $g(2400)$ to the refund attribute where the payoff difference is 2400. As a result—because $g(2400) > g(200)$ —she strictly prefers the refund. Intuitively, the single large payment of the refund draws the person’s focus more than the many small gains in monthly pay each month, leading her to overweight the refund.

In Section 3.1, we develop theoretical results that generalize this insight, and in Section 3.2 we show how our results help explain evidence in a number of domains. Section 3.3 discusses the relationship between our results and some implications of diminishing sensitivity identified in the literature.

3.1 Balanced and Unbalanced Choices

Bias toward concentration. We begin this section by formalizing and generalizing the example of tax refunds above. For any two alternatives c^i, c^j , let $A(c^i, c^j)$ be the set of attributes in which c^i is strictly better than c^j . For options c^1, c^2 , we say that c^1 *has more concentrated advantages than disadvantages relative to c^2* if $\min_{i \in A(c^1, c^2)} \Delta_i(\{c^1, c^2\}) > \max_{j \in A(c^2, c^1)} \Delta_j(\{c^1, c^2\})$ —that is, if c^1 ’s advantages relative to c^2 are uniformly greater than its disadvantages relative to c^2 . This condition holds for the tax example, because the first option’s advantage in the last payment (\$2,400) exceeds its disadvantage in the other 12 payments (\$200). Proposition 1 uses this condition to formalize the bias toward concentration:

Proposition 1 (Bias Toward Concentration). *Suppose that $g(\cdot)$ is continuous. Then, the following are equivalent:*

1. *Assumption 1 holds.*
2. *For any K and any c^1 and c^2 with the same consumption utility, if c^1 has more concentrated advantages than disadvantages relative to c^2 , then the agent strictly prefers to choose c^1 from*

$\{c^1, c^2\}$.

We also formalize a slightly different version of the bias toward concentration. For a third option c' in addition to c^1, c^2 , we say that c' combines two advantages of c^1 relative to c^2 if c' can be obtained from c^1 by merging c^1 's advantages in two attributes $k, l \in A(c^1, c^2)$ into a single attribute k . Formally, this holds if $u_k(c') - u_k(c^2) = [u_k(c^1) - u_k(c^2)] + [u_l(c^1) - u_l(c^2)]$, $u_l(c') = u_l(c^2)$, and $u_m(c') = u_m(c^1)$ for all $m \neq k, l$. We say that c' combines some advantages of c^1 relative to c^2 if there is a sequence $c(1) = c^1, c(2), \dots, c(n) = c'$ such that for each i , $c(i)$ combines two advantages of $c(i-1)$ relative to c^2 . We define when c' combines some disadvantages of c^1 relative to c^2 analogously.

Proposition 2 (Bias Toward Concentration). *Suppose Assumption 1 holds. If c' combines some advantages of c^1 relative to c^2 , and the agent is willing to choose c^1 from $\{c^1, c^2\}$, then she strictly prefers to choose c' from $\{c', c^2\}$. Conversely, if c' combines some disadvantages of c^1 relative to c^2 , and the agent is willing to choose c^2 from $\{c^1, c^2\}$, then she strictly prefers to choose c^2 from $\{c', c^2\}$.*

Proposition 2 says that concentrating an option's advantages makes it more preferable, while concentrating its disadvantages makes it less preferable. Intuitively, contrast a situation in which an option has advantages of 10 utils on each of two attributes with a situation in which it has an advantage of 20 utils on one attribute. In the former case, the difference of 10 utils on an attribute attracts some attention to that attribute and hence to that advantage, but it does not attract attention to the other 10 utils of advantage. In the latter case, however, each 10 utils of advantage attracts attention not only to itself, but also to the other 10 utils of advantage on the same attribute.

Balanced choices. To clarify the above intuitions as well as to explore further issues, we identify specific circumstances in which the bias toward concentration does not arise. As an example, suppose that the taxpayer above is asked not whether she prefers less withholding from her paycheck, but whether she wants a lump-sum payment of \$2,500 in December or a lump-sum payment of \$2,400 at tax time. This is a "balanced" choice because the advantages of the two alternatives are equally concentrated—on a single payment each. Using Equations (1) and (2), the person

strictly prefers to choose the first option if and only if $g(2500) \cdot 2500 > g(2400) \cdot 2400$. This holds if Assumption 1 holds, so that the consumer maximizes consumption utility. While applying the focus weights changes the agent’s perceived utility of the options, it does so in a way that reinforces the consumption-utility ordering.

To generalize this example, we say that a consideration set $C = \{c^1, c^2\}$ has *balanced tradeoffs* if for some K' we have $u_1(c^1) - u_1(c^2) = u_2(c^1) - u_2(c^2) = \dots = u_{K'}(c^1) - u_{K'}(c^2) \equiv a > 0$, $u_{K'+1}(c^2) - u_{K'+1}(c^1) = \dots = u_{2K'}(c^2) - u_{2K'}(c^1) \equiv b > 0$, and $u_k(c^1) = u_k(c^2)$ for $k > 2K'$. That is, relative to c^2 , c^1 is better by a on K' attributes and worse by b on K' attributes. Then, a consumption-utility maximizer chooses option 1 if $a > b$ and option 2 if $a < b$. Our agent does the same:

Proposition 3 (Rationality in Balanced Tradeoffs). *Suppose Assumption 1 holds. In a balanced tradeoff, the decisionmaker makes a consumption-utility-maximizing choice.*

Extending the logic of Proposition 3, the agent chooses rationally whenever the vector of utility differences in which c^1 is better than c^2 is a constant multiple of the vector of utility differences in which c^1 is worse than c^2 —that is, when the advantages of c^1 relative to c^2 have the same pattern across attributes as its disadvantages.

Small stakes yield rational choices. To conclude this section, we point out that if $g(0) > 0$, the decisionmaker maximizes consumption utility—and hence her bias toward concentrated advantages disappears—in the limit as stakes become small:

Proposition 4 (Rationality with Small Stakes). *Suppose $g(0) > 0$ and $g(\cdot)$ is continuous at zero. For any $u = (u_1, \dots, u_K)$ and $\delta = (\delta_1, \dots, \delta_K)$, there is an $\bar{\epsilon} > 0$ such that if $0 < \epsilon < \bar{\epsilon}$, the decisionmaker makes a consumption-utility-maximizing choice from the choice set $\{u, u + \epsilon \cdot \delta\}$.*

Intuitively, with small stakes, the focus weights on all attributes are near their minimal level $g(0)$, so that if $g(0) > 0$, focus on all attributes is about equal. Both because small-stakes decisions are all else equal less interesting and because we do not know whether $g(0) > 0$ holds in general, Proposition 4 is economically the least important of our predictions. But the result does highlight an interesting way our model is different from theories of attention, such as Reis (2006), Gabaix,

Laibson, Moloche and Weinberg (2006), Sims (2010), and Salant (forthcoming), in which the agent makes rational choices taking into account the cognitive or effort cost associated with attention or calculation. In these models, an increase in stakes increases the value of exerting cognitive effort, and because a rational agent responds to this incentive, on average she makes better decisions for larger stakes. In our model, in contrast, focusing distorts behavior in non-trivial unbalanced decisions, while rationality may obtain with small stakes.⁶

3.2 Evidence on the Bias Toward Concentration

In this section, we discuss a variety of evidence consistent with our prediction of a bias toward concentration. In light of the equivalence result in Proposition 1, much of this evidence can be interpreted as directly supporting Assumption 1.

Like much of the economic evidence on salience, most of the facts we review here concern attitudes toward monetary payments. To explain this evidence using our model, we will assume that individuals separate, in specific ways, different components of monetary transactions into different attributes. The hypothesis that people separate money into different “mental accounts” of wealth, income, and expenditures has long been considered a stylized fact, and forms the basis of Thaler’s (1985, 1999) theory of mental accounting. Temporal separation, in particular, facilitates the separation of monetary transactions in a person’s mind (Thaler and Johnson 1990). Unfortunately, however, there is little evidence on exactly how individuals compartmentalize money. Here we make what we believe are natural assumptions, but it is important to emphasize that our analysis does depend on these assumptions.

Concentrated payments. In several economic settings, individuals choose one-time payments over streams of payments, thus exhibiting behavior consistent with a bias toward concentrated advantages. In the tax refund context which motivated our our example at the beginning of the section, Jones (2010) documents that nearly 80% of taxpayers receive a lump-sum tax refund when filing their taxes, even though they could receive the same nominal amount in the form of

⁶ If the function $g(\Delta)$ has an upper bound, our model also implies close to consumption-utility-maximizing behavior if stakes are sufficiently large in *all* attributes. Even then, the agent exhibits a bias toward concentration when differences are smaller in at least some attributes.

higher monthly paychecks the previous year. In 2004, refunds totaled \$155 billion, and comprised on average 7 percent of adjusted gross income for refund recipients. Furthermore, it is unclear whether previous explanations based on the penalties for tax underpayment (Highfill, Thorson and Weber 1998), a commitment to save (Thaler 1994), and inertia (Jones 2010) explain the very high prevalence of refunds. Our model says that taxpayers might like the tax refund simply because it comes as a single attention-grabbing big payment rather than as a sequence of less-noticed small payments.

Similarly, in retirement planning, retirees seem to take too much of their retirement wealth in one lump sum rather than as an annuity.⁷ As a manifestation of this phenomenon, 59% of respondents in the Health and Retirement Survey report that they would accept \$500 less in Social Security benefits in exchange for a one-time lump-sum payment of \$87,000, where the latter sum was chosen to be actuarially fair for the average person (Brown, Casey and Mitchell 2008).⁸ Once again, our model provides a very simple explanation for this example and the “annuity puzzle” more generally: retirees simply like a lump-sum payment because it looks very large relative to an annuity’s monthly payments.⁹

Product design and pricing. Because having to pay is a disadvantage of buying a product, Proposition 2 implies that splitting a product’s price into multiple components increases demand for it. This can help explain the popularity of such pricing attempts by firms.¹⁰ As a common example, most retailers post pre-tax prices for products, which can be seen as separating the tax into a different price component. Indeed, Chetty et al. (2009) document that—consistent with

⁷ For long-standing theoretical arguments that risk-averse individuals should take much of their retirement income in the form of an annuity, see for example Yaari (1965) and Davidoff, Brown and Diamond (2005). For a review summarizing evidence and arguments that current annuitization levels are too low even taking into account adverse selection and other classical considerations, see Brown (2009).

⁸ For couples, the actuarially fair lump-sum payment was \$80,000, which is slightly lower than that for individuals because Social Security payments can decrease when either spouse dies. Hence, for couples the question posed involved an \$80,000 lump-sum payment.

⁹ Johnson, Hershey, Meszaros and Kunreuther (1993) also find a preference for concentrated payments in the context of insurance. In a hypothetical choice between two five-year disability policies, the majority of subjects prefer the one with a \$20 higher monthly premium and \$1,200 refund if no disability claim is made. Abstracting away from the possibility that a disability claim might be made, this is exactly what Proposition 1 says. Allowing for disability to occur makes our point stronger, because the refund is conditioned on no disability but the monthly premium payments up to the disability event are not.

¹⁰ We explore the implications of the bias toward concentration for firm pricing in more detail in Section 5.

our model—demand for a taxable item is lower if the tax is included in the posted price than if it is added at the register. In a more specific setting, Woodward and Hall (2010) argue that the underappreciation of mortgage-broker compensation they document in borrowers arises partly because such compensation is artificially split into many fees. Similarly, many retailers and service providers, including credit-card issuers, banks, mobile-phone companies, and airlines, make a large part of their profits from imposing a multitude of relatively small fees that may not seem like much to consumers when getting the product, but that can easily add up to significant amounts.¹¹

Our result that price splitting increases demand may also contribute to the frequency of financing for consumer purchases: a laptop which costs “\$29 a month” may seem less expensive than one with a price of \$899. Because our model does not rely on liquidity constraints, it also explains why consumers sometimes resort to financing even when they have liquid funds available (Bertaut, Haliassos and Reiter 2009, Stango and Zinman 2009).¹² Also consistent with the psychological mechanism, although not with the precise formal implementation of our theory, are “pennies-a-day” marketing strategies, in which prices such as magazine or newspaper subscription fees are quoted in disaggregated (monthly or daily) terms even when payments remain aggregated.¹³

Finally, although the reasons given are varied and sometimes not clearly related to focusing, the bias toward concentration is reminiscent of views in the marketing literature that successful product design and placement concentrates on one or a few core product attributes. For example, Anderson et al. (2006) argue that a firm’s “value proposition”—the firm’s statement that summarizes why a consumer should buy its product—is effective if it demonstrates superiority on a few elements, even if this means ceding a bit on other elements. Expressing a similar view in the context of brand positioning, Aaker (1991) writes that “[p]robably the most used positioning strategy is to associate an object with a product attribute or characteristic” (page 114), and argues that associating a

¹¹ For instance, Agarwal et al. (2007, 2008) and Stango and Zinman (2009) document that consumers pay very significant amounts in banking and credit-card fees, and the Federal Communication Commission is considering regulating the mobile-phone market to help consumers avoid the “bill shock” from unexpectedly large bills due to many minor charges (Federal Communication Commission 2010).

¹² Stango and Zinman (2011) propose a “fuzzy math” model of borrowing which also predicts that quoting the price of a loan in terms of monthly payments makes the loan seem cheaper to borrowers. The mechanism driving this effect, however, is completely different from ours, and only holds for short-term loans.

¹³ Another prominent example is when charities claim that donors can do good with pennies a day. Gourville (1998) presents a systematic study showing the effectiveness of this strategy.

product with many characteristics is counterproductive.

3.3 Bias Toward Concentration and Diminishing Sensitivity

Before moving on to applications, we briefly discuss the connection between our model and diminishing sensitivity, the basic hypothesis of prospect theory (Kahneman and Tversky 1979) that an initial deviation from a reference point is felt more strongly than an increase in the deviation. As emphasized by Thaler and Johnson (1990) and others, diminishing sensitivity predicts that individuals should prefer to segregate gains and integrate losses. This prediction is in direct contradiction with our model’s prediction of a bias toward concentration, and most notably with Proposition 2.

The most compelling evidence for diminishing sensitivity comes from comparing people’s attitudes toward changes starting from outcomes closer to versus further from the reference point.¹⁴ Combined with diminishing sensitivity in consumption utility within each attribute, our model would accommodate all this evidence. Within the combined framework, diminishing sensitivity’s prediction about separating gains and combining losses would be weakened or reversed by the bias toward concentration, explaining why there is limited evidence for this prediction.¹⁵ More interestingly, our theory seems useful for making sense of some examples commonly used to illustrate diminishing sensitivity in the literature. One popular example argues that consumers are more willing to buy add-ons such as a car stereo after buying the car itself, due to diminished sensitivity to additional expenditures following a big payment (Thaler 1985, Chiu and Wu 2009, for example). This example actually *contradicts* diminishing sensitivity: due to diminishing sensitivity on the “car” attribute, the consumer should be more willing to make an unrelated purchase. In contrast, our model explains why the two purchases must be related: a consumer is more likely to purchase a car stereo because the high-value car tilts her focus toward the “car” attribute, which increases the perceived value of the add-on as well.

¹⁴ For instance, people are risk-averse over gains and risk-loving over losses (Kahneman and Tversky 1979), people are more willing to work to save a given amount when it comes off a small payment rather than a large payment (Tversky and Kahneman’s (1981) famous calculator example), and—gambling that the price will bounce back—people are more likely to hold on to losing assets than to winning assets (Odean 1998, Genesove and Mayer 2001).

¹⁵ Thaler and Johnson (1990) and Linville and Fischer (1991), find that subjects do prefer to separate gains, but—in contrast to the prospect-theory prediction—they also prefer to separate losses. Lehenkari (2009) documents that individual investors in the Finnish stock market do not integrate sales of stocks that have lost money or segregate sales of stocks that have made money.

4 Intertemporal Choice

This section explores some implications of our model for choice over time. A fundamental insight is that how a person weights an action’s costs and benefits depends not on their temporal placement, but on their *concentration*. In addition, because the same dynamic choice situation can generate different concentrations of costs and benefits when considering a single period’s choice than when considering the entire problem, our model predicts a form of time inconsistency in behavior.

By endogenizing time preference and time inconsistency based on the concentration of the costs and benefits of available actions, our theory contributes to the literature in two main ways. First, by predicting how the emergence of present-biased versus “future-biased” behavior depends on the economic environment, our theory provides a new perspective on several distinct phenomena discussed in the literature. Second, in contrast to the existing work on time-inconsistent behavior, our theory provides an unambiguous welfare measure without a-priori assuming that the ex-ante choice is welfare relevant, and identifies a simple way to determine whether ex-ante or ex-post choices better reflect welfare.

In Section 4.1, we develop the key ideas of this section through a stylized model of an investment decision, and discuss how our results relate to previous findings. In Section 4.2, we identify general conditions for when different forms of time inconsistency occur.

4.1 A Simple Investment Decision

We develop the central ideas of this section using a simple dynamic decision problem that can be interpreted as a stylized model of many intertemporal decisions, including exercise, work, harmful consumption, and consumption-savings. There are $T_i + T_b$ periods, $t = 1, \dots, T_i + T_b$. The agent makes an investment or effort decision $e_t \in \{0, 1\}$ in each of the the first $T_i \geq 1$ periods, and her investments provide benefits in the next $T_b \geq 1$ periods. The consumption-utility cost of investment in each of the first T_i periods is $e_t \cdot B$, and the consumption-utility benefit in each of the last T_b periods is $u_t = A(\sum e_t \cdot B) / T_b$. The variable A measures the efficiency of the investment: investing

maximizes consumption utility if and only if $A \geq 1$ holds.¹⁶

Our simple model abstracts away from many aspects of the real-life settings motivating it, including possible overlaps between the investment and benefit periods and non-binary choices. Adding these considerations to our framework affects the focus weights and hence the precise predictions, but does not seem to eliminate the basic intuitions we identify. In addition, Section 4.2 develops general versions of our most important predictions about time-inconsistent behavior.

Proposition 5 below characterizes ex-ante and ex-post choices in the above choice problem, using the intertemporal choice model developed in Section 2.2. Recall from Section 2.2 that in the ex-ante or commitment choice problem the agent makes decision at $t = 1$ about her action in every period, while in the ex-post or no-commitment choice she makes daily decisions given her beliefs about future behavior.¹⁷ Before we state the formal results, we illustrate them using two economically relevant examples.

Example 1 (Exercise). A consumer is deciding in each of the periods 1 through 100 whether to exercise. Exercising in a period generates pain of 80 in that period and health benefits of 1 in each period 101 through 200. ($T_i = T_b = 100, B = 80, A = 5/4$.)

Because $A > 1$, the consumption-utility-maximizing choice is to exercise. Without commitment, the consumer's decision of whether to exercise on any given day can be represented as a choice between $(-80, 1, \dots, 1)$ and $(0, \dots, 0)$, where 80 is the pain from exercising today, and the 1's represent the health benefits in the 100 future periods. Hence, if $g(80)/g(1) > 100/80 = 1.25$ —a condition that may well hold given that $g(\cdot)$ is increasing—the consumer never exercises. Because exercising has an attention-grabbing concentrated current cost and easy-to-neglect dispersed future benefits, the consumer is biased toward the present.

Now consider what the consumer would commit to in Example 1. Due to the linearity and symmetry of the problem, her choice is effectively between always and never exercising, $(-80, \dots, -80, 100, \dots, 100)$ and $(0, \dots, 0)$. Because each alternative is better than the other on exactly 100

¹⁶ The parameter B scales both the costs and benefits of effort; its main role is to guarantee that our simple numerical examples map to the formal setup.

¹⁷ Throughout this section we assume that the agent's beliefs about future behavior are correct. It is easy to check that in the current setting, naive beliefs generate the same behavior.

attributes ($T_i = T_b$), this is a balanced choice, so that the consumer commits to the consumption-utility-maximizing option of always exercising. Intuitively, looking at exercising from the perspective of a lifetime rather than a single day leads the consumer to integrate the incremental health benefits of exercise, leading her to focus more on these benefits.

Our next example considers choices when investment is directed toward a concentrated goal.

Example 2 (Writing a book). In each of the first 100 periods, a scholar can expend effort to write a book, whose benefit is realized in period 101. A period's effort has a utility cost of 50 in that period and increases the utility from the book by 40. ($T_i = 100, T_b = 1, B = 50, A = 0.8$.)

Because $A < 1$, the consumption-utility-maximizing choice is never to write the book. Without commitment, the scholar's choice in each period is equivalent to that between $(-50, 40)$ and $(0, 0)$, trading off the current pain of writing with an incremental improvement in the book. Since this is a balanced choice, the scholar makes the consumption-utility-maximizing choice of never working on the book.

From an ex-ante perspective, however, the choice of whether to write the book is effectively between $(-50, \dots, -50, 4000)$ and $(0, \dots, 0)$, so that if $g(4000)/g(50) > 1.25$, the scholar commits to writing the book. Intuitively, in her ex-ante choice the scholar focuses on the large concentrated gain from having a completed book and pays relatively less attention to the dispersed everyday costs, so that she is now biased toward the future.

We now turn to stating our general results on the agent's behavior and biases. Notice that because of linearity and symmetry, in both the ex-ante and ex-post problems the agent prefers to exert effort either in all periods or in no period, and is indifferent only in a knife-edge case. Let A_{ante}^* and A_{post}^* be the cutoff levels of A above which the agent chooses to exert effort in the ex-ante and ex-post problems, respectively. For each $i = ante, post$, the decisionmaker is biased toward the present in ex- i choice if $A_i^* > 1$ (because to invest she requires the investment to be more efficient than a consumption-utility maximizer would), and she is biased toward the future if $A_i^* < 1$.

Proposition 5. *Suppose $g(\cdot)$ satisfies Assumption 1, and is continuous. Then:*

1. *The cutoff A_{post}^* does not depend on T_i , is strictly increasing in T_b , and equals 1 if $T_b = 1$.*

2. The cutoff A_{ante}^* is strictly increasing T_b/T_i , and equals 1 if $T_b = T_i$.
3. If $T_i = 1$, $A_{ante}^* = A_{post}^*$, and if $T_i > 1$, $A_{ante}^* < A_{post}^*$.

Part 1 states that the agent is ex post biased towards the present relative to her welfare-maximizing choice whenever $T_b > 1$, and more so the higher T_b . Intuitively, when benefits are distributed over multiple periods like in the exercise example, the agent focuses less on them and hence overweights the cost of effort. Part 2 shows that in her commitment choice, the agent is more biased towards the future the more concentrated benefits are (T_b low) and the more dispersed costs are (T_i high). Rational choice obtains when $T_i = T_b$, like in the exercise example, while the agent overcommits to the future when $T_i > T_b$, like in the book example.

In both the exercise and book examples, the consumer's ex-post behavior is less future-oriented than her ex-ante behavior. Part 3 of the Proposition says that the same happens for any $T_i > 1$. Intuitively, since the per-period benefits from multiple investments accumulate, the range of possible future benefits is larger in the broader perspective of the ex-ante decision than in the narrower perspective of the ex-post decision. Because the range of investment costs is the same, the agent focuses relatively more on the future benefits ex ante than ex post.

Analogously, our model also predicts intuitively plausible circumstances in which the agent is *less* present-oriented ex post than ex ante, when the range of investment costs is narrower ex post than ex ante.

Example 3 (Adding Achievements). Consider the decisionmaker from Example 2, and suppose that in period 91, she can decide whether to commit to writing an editorial for the newspaper. Writing has an effort cost of 5 for the next 10 periods, and a benefit of 45 in period 102.

Because $45 < 10 \cdot 5$, not writing the editorial is the consumption-utility-maximizing choice. When looking at the book and editorial-writing decisions together ex ante, the range of possible effort levels in periods $t = 91, \dots, 100$ is 0 to 55, so the effort dimension looms large relative to the benefit of the editorial. Hence, while the scholar commits to writing the book, she would prefer not to take on the editorial assignment. Now suppose that she has committed to writing the book, and makes the decision regarding the editorial in period 91. At this point, whether to take on the

new assignment is equivalent to choosing between $(-50, -50, \dots, -50, 0)$ and $(-55, \dots, -55, 45)$. Since the *incremental* per-period costs of the editorial seem small relative to the large benefit, the scholar might commit to writing editorial as well.

In the rest of this subsection, we discuss some economic implications of the results above.

Lifestyle choices versus achievements. Proposition 5 predicts a time-inconsistent underinvestment into the future when $T_i \approx T_b \gg 0$, and a time-inconsistent overcommitment to investing in the future when $T_b \ll T_i$. The former type of situation matches a number of lifestyle choices, including consumption-savings and borrowing behavior (Laibson 1997, Laibson, Repetto and Tobacman 2007, Shui and Ausubel 2004, Skiba and Tobacman 2008, Meier and Sprenger 2010, for example), the consumption of harmful products (O’Donoghue and Rabin 1999, Gruber and Kőszegi 2001, Gruber and Kőszegi 2004, O’Donoghue and Rabin 2006, and others), and exercise (DellaVigna and Malmendier 2004, DellaVigna and Malmendier 2006, Acland and Levy 2010), which the literature on present bias has invoked as prime examples of situations in which people act in an overly present-oriented way. The latter prediction is consistent with the view of several researchers, including Scitovsky (1976), Loewenstein, O’Donoghue and Rabin (2003), Kahneman et al. (2006), and Hamermesh and Slemrod (2008), that many people choose overly ambitious careers relative to what would make them happiest. Most closely related to the mechanism in our model, Kahneman et al. (2006) argue that people are too keen to pursue career advancement because they overweight the big potential achievements, such as a big raise or a promotion, relative to the many everyday costs.

Welfare. The existing literature on present bias has not converged on an entirely satisfactory view of the welfare ranking of different choices. One approach is to assume, based on introspection, that a person’s high relative weight on the present is uniformly mistaken or irrelevant, and evaluate welfare using long-run preferences. Another approach is to use the Pareto criterion as applied to the intertemporal incarnations of the individual, which, however, is often silent about the welfare-optimal choice.¹⁸ Going beyond both of these approaches, our model yields an unambiguous welfare

¹⁸ For examples of the former view, see for instances DellaVigna and Malmendier (2004), Gruber and Kőszegi (2004), and O’Donoghue and Rabin (2006); for examples of the latter view, see for example Laibson (1997), Bernheim and Rangel (2008), and Asheim (2008).

ranking in all situations without a-priori assuming that the ex-ante perspective is the welfare-relevant one. The guiding principle emerging from our theory is that decisions in more balanced choices are more likely to reflect a person’s true well-being. To illustrate, recall Examples 1 and 2 on exercising and writing a book, respectively. In both of these examples, the agent is time-inconsistent in that her ex-post choices are more present-oriented than her ex-ante choices. But—corresponding to which choice is balanced in each case—in the exercise example the agent’s welfare is aligned with her ex-ante preferences, while in the book-writing example it is aligned with her ex-post preferences.¹⁹ More generally, Part 2 of Proposition 5 says that the ex-ante welfare criterion is appropriate if $T_i = T_b$ —a condition approximately satisfied for lifestyle choices—and Part 1 says that the ex-post criterion is appropriate if $T_b = 1$ —a condition often approximately satisfied for achievement-type tasks.

New comparative statics. Our model also makes a number of new comparative-statics predictions on the nature and extent of biases in intertemporal choice, which seem intuitively compelling and provide testable out-of-sample predictions for our theory.

Part 1 of Proposition 5 says that present bias in ex-post choice (relative to the utility-maximizing outcome) is increasing in the number of periods T_b in which the consequences of current misbehavior are dispersed. Intuitively, an increase in T_b dilutes the benefits but not the cost of exerting effort today, leading the agent to focus relatively less on the benefits. This prediction distinguishes our model from hyperbolic discounting. Suppose, for instance, that an employee must decide whether to perform a five-hour job today or put it off to the future. Our model says that she is more likely to perform the job if the alternative is a seven-hour job in six weeks than if the alternative is 10 minutes of extra work each day for the next six weeks. Hyperbolic discounting says that this might not make much difference.²⁰

Proposition 5 also implies that the degree of time inconsistency—the difference in the agent’s ex-ante and ex-post preferences—is increasing in the number of investment periods T_i . Intuitively,

¹⁹ The ex-ante approach above would by assumption impose in both cases that the person’s ex-ante choice is optimal. Because ex-ante and ex-post behavior differs, the Pareto approach would not make a welfare ranking in either example.

²⁰ In fact, because the bulk of work is sooner in the latter case, hyperbolic discounting might typically predict that the employee is more likely to do the five-hour job if the alternative is 10 minutes per day for six weeks.

because T_i determines how many benefits the agent integrates in the ex-ante choice that she considers one-by-one in the ex-post choices, it determines the extent to which the ex-ante and ex-post perspectives differ. This logic is different from that in models of present bias—where time inconsistency arises because the relative weight attached to utility in a period changes once a person is in that period—and yields yet more predictions that distinguish our model from hyperbolic discounting. Most importantly, a person is more likely to commit for a substantial part of her future rather than only a trivial part of her future—because with the former type of commitment she is more likely to feel that she has a noticeable impact on her life. Although we are not aware of systematic tests of this prediction, it seems consistent with the observation that despite apparent present bias, the take-up of even effective short-term commitment devices in the field has been quite low (see Bryan et al. 2010 for a review).²¹

We mention two other new predictions briefly. First, Part 2 of Proposition 5 implies that the more dispersed are an investment’s costs and the more concentrated are its benefits, the greater the agent’s tendency to overcommit ex ante. For example, people often commit to months of hard work for a salary increase, but do not commit to leaving their family vacation one day early for a \$100 one-time job. Second, when there are multiple possible investments to be made, the agent may agree to more utility-decreasing future-oriented tasks as time goes on. Consistent with this prediction, many people—including academics—appear to pile on tasks, only to bitterly complain later that this has made them worse off.

4.2 General Conditions for Time-Inconsistent Behavior

We conclude this section by showing that our central insights about time-inconsistent behavior extend to more general settings. To do so, we introduce some notation and concepts. Suppose there are κ attributes at time t , and let the corresponding utility functions be $u_{t1}(\cdot)$ through

²¹ Because our model assumes time-consistent consumption utility, unlike present bias it also predicts that the decisionmaker’s commitment and no-commitment choices over any *single decision* are identical—because the identical choice sets generate the same focus. For example, a person who would strongly prefer to commit to exercising over her entire life may—if this was the only choice she could now make—commit to *not exercising* on June 16, 2019. Because we believe that individuals also have discounting-based time inconsistency, we expect that individuals often have time-inconsistent preferences even for single decisions. Rather than this extreme prediction on single choices, we view the above comparative static as the more important prediction of our model.

$u_{t\kappa}(\cdot)$, so that $u_t = \sum_{\kappa=1}^{\kappa} u_{t\kappa}$. We denote the history of consumption choices up to period t by $\bar{c}_t = \{c_t, \dots, c_1\}$. We allow $u_{t\kappa}(c_t, \bar{c}_{t-1})$ to depend on both current and past consumption choices, but require that the choice set C_t does not depend on previous choices.

To explore the question of time inconsistency, we identify the types of choices the agent prefers if she is allowed to deviate from her ex-ante committed plan in period t , assuming that in all other past and future periods she is still committed to the plan. Formally, for an ex-ante committed plan (c_1^*, \dots, c_T^*) , we ask how, given history c_1^*, \dots, c_{t-1}^* , the decisionmaker's choice in the continuation problem $C_t \times \{c_{t+1}^*\}, \dots, \times \{c_T^*\}$ compares to c_t^* .

Our first result identifies a general set of circumstances under which the agent prefers to behave in a more present-oriented way ex post than she would have committed to ex ante. We say that the current utility of a current action is not controlled by previous choices if, whenever $u_{t\kappa}$ depends on c_t , it does not depend on the history \bar{c}_{t-1} . We say that period- t choice has a unidirectional effect on future utility if for any $c_t, c'_t \in C_t$, either $u_{t\kappa}(c'_t, \bar{c}_{t-1}) \geq u_{t\kappa}(c_t, \bar{c}_{t-1})$ for all κ and $u_{\tau\kappa}(\bar{c}_{\tau,-t}, c'_t) \leq u_{\tau\kappa}(\bar{c}_{\tau,-t}, c_t)$ for all $\tau > t, \kappa$, and feasible c_{t+1} through c_T , or vice versa. Both of these conditions are satisfied for example by the investment problem of Section 4.1.

Proposition 6 (Present-Oriented Time Inconsistency). *Suppose $g(\cdot)$ satisfies Assumption 1. Take any $t \in \{2, \dots, T\}$, and suppose that (i) the current utility of current actions in period t is not controlled by previous choices; and (ii) period- t choice has a unidirectional effect on future utility. Then, the decisionmaker would prefer an ex-post choice in period t which yields at least as high period- t utility as her ex-ante plan.*

Because of (i), the decisionmaker focuses on the period- t utility consequences of c_t equally ex ante and when she makes her choice in period t . Because she has (weakly) fewer choices in period t than ex ante, however, her focus on future periods $\tau > t$ is (weakly) smaller at time t than ex ante. In addition, by (ii) the agent's her choice at time t is a clear trade-off between the present and the future. Since she pays (weakly) less attention to the future in relative terms at time t than ex ante, it follows that she will act in a (weakly) more present-oriented manner.

We next identify conditions under which the decisionmaker can be prone to time inconsistency in future-oriented behavior, whereby she keeps becoming more and more ambitious in her quest for

achievements. We say that a future attribute τ^*, k^* is fully controlled by period- t choice if $u_{\tau^*k^*}$ depends only on c_t . This attribute is said to have a unidirectional effect on the utility of other attributes if, for any c_t and c'_t such that $u_{\tau^*k^*}(c'_t) > u_{\tau^*k^*}(c_t)$, for any other τ, k with $\tau \geq t$ it must be that $u_{\tau,k}(c_\tau, c_{\tau-1}, \dots, c_{t+1}, c'_t, c_{t-1}, \dots, c_1) \leq u_{\tau,k}(c_\tau, c_{\tau-1}, \dots, c_{t+1}, c_t, c_{t-1}, \dots, c_1)$.

Proposition 7 (Future-Oriented Time Inconsistency). *Suppose $g(\cdot)$ satisfies Assumption 1. Consider any $t \in \{1, \dots, T-1\}$, and suppose that (i) there is a future attribute (τ^*, k^*) fully controlled by period- t choice, and (ii) this attribute has a unidirectional effect on the utility of other attributes. Then, the decisionmaker would prefer an ex-post choice in period t which yields at least as high utility in attribute (τ^*, k^*) as her ex-ante plan.*

This result generalizes Example 3. Thinking of the future utility dimension (τ^*, k^*) as an achievement, the statement means that people are (weakly) too prone to take on an additional task than they would have preferred earlier. Intuitively, because of (i) the range of utility from the achievement is the same ex ante and ex post, while the range of costs is (weakly) smaller ex post. Hence, the benefit from the achievement looms (weakly) larger in relative terms ex post than ex ante. Since by (ii) the individual faces a clean trade-off between the benefit and the costs, she is (weakly) more willing to take on the additional task ex post.

5 Product Design and Pricing

In this section, we explore some implications of our model for profit-maximizing firms' product design and pricing behavior. The idea that firms design and position products in part to attract consumer attention is central to both academic and applied marketing research. By specifying how product features determine consumer focus, our model provides one way to incorporate this key marketing idea into economics. Although a full analysis of the issues we raise is beyond the scope of this paper, we formalize and qualify a number of insights from marketing that have not been explored in economics, and derive new predictions. Our analysis complements research by Eliaz and Spiegler (2011a, 2011b) on the competitive implications of marketing, which assumes that firms can, at an exogenously given cost, manipulate consumers' consideration sets or offer attention-grabbing

products. Because our model is based on a fully articulated theory of consumer behavior, we can make specific predictions on the product features firms will use to influence consumers' focus.

5.1 Product Design Under Monopoly

We first explore how a monopolist designs and prices products to take advantage of consumers' bias toward concentration. Suppose that a monopolist can sell exactly one item of one product to a consumer with focus-dependent behavior. The product is a vector $(v_1, \dots, v_{K_1}, -p_1, \dots, -p_{K_2})$, where the first K_1 dimensions measure consumption value over different product attributes, and the last K_2 dimensions are price (or other cost) attributes to the consumer. We assume that consumption utility over all attributes is linear, and that the outside option of the consumer is not purchasing the product, which is represented by the vector $(0, \dots, 0)$. The monopolist chooses $v_i, p_j \geq 0$ to maximize

$$\sum_{i=1}^{K_2} p_i - \phi \left(\sum_{i=1}^{K_1} v_i \right).$$

Hence, the firm can produce each of the value dimensions at the same marginal cost, and cares about the total price the consumer pays. We assume that the cost function $\phi(\cdot)$ is twice continuously differentiable and satisfies $\phi'(0) = 0, \phi''(c) > 0$ for all $c \geq 0$, as well as $\lim_{c \rightarrow \infty} \phi'(c) = \infty$.

Before turning to results, we define, for any utility difference Δ , the attention-weighted difference $G(\Delta) \equiv g(\Delta)\Delta$. Under Assumption 1, $G(\Delta)$ is strictly increasing and convex near zero. For Proposition 8 below, we make the natural assumption that the convexity of $G(\Delta)$ extends to the relevant range of medium stakes. Intuitively, this assumption means that adding to a larger advantage of a product affects a consumer's perceived utility more than adding to a smaller advantage, as she is focusing more on the attribute with the larger advantage.²²

Proposition 8 identifies how the monopolist responds to the bias toward concentration:

²² When the available options differ more rather than less in a given attribute, how a further increase in the differences in this attribute affects the difference in perceived utility is determined by three effects. First, since focus on the attribute is greater, the increase in the difference counts more heavily in perceived utility. Second, the increase in attention applies to a greater utility difference, again increasing the difference in perceived utility. Third, depending on the shape of $g(\cdot)$, the increase in attention may be smaller, decreasing the effect on perceived utility. The assumption that $G(\cdot)$ is convex over a range amounts to assuming that in as much as there is such a negative effect, the former effects outweigh it. This is especially plausible since for moderate-range stakes, one might expect the first direct effect to dominate.

Proposition 8. *Suppose $g(\cdot)$ satisfies Assumption 1 and is differentiable, and $G(\Delta)$ is convex. The optimal product has $v_i > 0$ for one value attribute and $v_j = 0$ for all other $j \neq i$ value attributes; and has $p_1 = p_2 = \dots = p_{K_2} > 0$. If $K_2 = 1$, the level of value production is efficient ($\phi'(v) = 1$), and if $K_2 > 1$, there is overproduction of value ($\phi'(v) > 1$).*

Consistent with the evidence and discussion in Section 3.2, the firm concentrates the value of the product into a single attribute to attract focus to this value, while it evenly distributes the price among components to minimize the focus on the price. To understand the efficiency implications, first note that with a single price dimension ($K_2 = 1$), the consumer faces a balanced choice, and because she acts rationally in these situations, the monopolist has an incentive to choose the efficient quality level. With at least two price dimensions, however, the consumer focuses more on the single large value than on the many small prices, and in response the monopolist produces “too much” value.²³

5.2 Competition and Relative Positioning

We next explore incentives for relative positioning and value production under competition. One positioning strategy firms often seem to attempt in practice is to build products with a unique attribute in which other firms are perceived to be significantly weaker. For example, Volvo has sought leadership in safety and BMW in performance. This strategy echoes the result of Proposition 8 that a firm should concentrate on one distinguishing value attribute, and is consistent with common advice on brand positioning. At the same time, some firms seem to attempt the exact opposite strategy—they try to copy the identifying attribute of the competitor. For instance, since the introduction of the iPhone, touch-screen mobile phones have become ubiquitous. This leads to the natural question: when should a firm copy the flagship attribute of its competitor, and when should it differentiate by emphasizing an attribute in which the competition is weak?

To address the above question, we consider a simple model of how a firm responds to a competitor’s exogenously given product. We assume that products have two quality attributes ($K_1 = 2$)

²³ While the sharp form of the results in Proposition 8 relies on our assumptions that attributes are symmetric and perfectly substitutable in production and consumption, the basic mechanism of concentrating value and separating prices to bias focus toward buying the product is independent of these assumptions.

and a single price dimension ($K_2 = 1$). Firm 1 chooses v_1, v_2, p , where v_1 and v_2 are the qualities on the two attributes and p is the price. The total cost of production is $\phi(v_1 + v_2)$. Firm 2 sells an exogenously given competing product with qualities \bar{v}_1 and $\bar{v}_2 < \bar{v}_1$ and price \bar{p} . The consumer's consideration set includes these two products, as well as the option of buying neither product.²⁴

Proposition 9 characterizes firm 1's optimal response depending on how its product's quality compares to the overall quality of firm 2's product.

Proposition 9. *Suppose $g(\cdot)$ satisfies Assumption 1 and is differentiable.*

(1) *If $v_i < \bar{v}_i$ for $i = 1, 2$, then $v_1 > v_2$.*

(2) *If $v_1 + v_2 > \bar{v}_1 + \bar{v}_2$ and either $g(v_1 + v_2)$ is large enough or \bar{v}_2 is small enough, then $v_1 < v_2$.*

Proposition 9 provides a simple explanation for why firms sometimes position products with respect to competitors' weaknesses and sometimes copy competitors' strengths. Part (1) says that when firm 1's product is dominated by firm 2's product, then firm 1 chooses as its stronger attribute firm 2's stronger attribute. In contrast, Part (2) says that if firm 1 is sufficiently better than firm 2, then firm 1 chooses as its stronger attribute firm 2's weaker attribute. To understand the intuition, suppose that firm 2 only produces on attribute 1 ($\bar{v}_2 = 0$), and consider firm 1's choice of whether to allocate a given value v to attribute 1 or attribute 2. If $v < \bar{v}_1$, then allocating to attribute 1 is optimal: it maximizes the focus on firm 1's product without affecting the focus on the competitor's product. If $v > \bar{v}_1$, however, then allocating the value v to attribute 2 becomes optimal: by lowering the weight on attribute 1 from $g(v)$ to $g(\bar{v}_1)$, this reduces focus on the competitor's strength.²⁵

Having analyzed a firm's incentive to differentiate its product from those of competitors, to conclude this section we explore the resulting quality levels in "attentionally differentiated" and "attentionally undifferentiated" competition. In the former case, $K_1 = 2$, and firm $i = 1, 2$ can only

²⁴ We analyze firm 1's best-response function rather than a full equilibrium in which firm 2 also chooses strategically to avoid dealing with mixed-strategy equilibria. Hence, our analysis can be viewed as identifying what firms would ideally like to do in response to competitors' behavior, and how these strategic incentives play out in the market depends on the precise market environment and technologies available to firms.

²⁵ The additional assumptions in Part 2 of the proposition address the possibility that firm 2's value on attribute 2 is positive ($\bar{v}_2 > 0$), but "just under the radar" of consumer attention. In this case, switching to that attribute draws attention to firm 2 as well. However, when \bar{v}_2 is close to zero this effect is small, while when $g(v_1 + v_2)$ is large the effect in the text dominates.

produce attribute i . In the latter case, there is a single value dimension in which firms compete ($K_1 = 1$).²⁶ In both cases, we assume that there is only one payment dimension ($K_2 = 1$), and that the consumer’s consideration set includes both options as well as not buying. Firms simultaneously choose quality levels and price. Then:

Proposition 10. *Suppose $g(\cdot)$ satisfies Assumption 1 and is differentiable. For any symmetric pure-strategy equilibria, value is higher with attentionally differentiated than with attentionally undifferentiated competition, and in either case higher than under monopoly.*

The difference between differentiated and undifferentiated competition arises from how an increase in a firm’s quality affects the consumer’s focus. With differentiated competition, an increase in firm 1’s quality draws attention to attribute 1, benefitting firm 1 and hurting firm 2. With undifferentiated competition, however, an increase in firm 1’s quality either has no effect on the consumer’s attention (if firm 2’s product is superior), or it draws more attention to quality (if firm 1’s product is superior), but in both cases it has the same effect on the two firms. As a result, firm 1 has a greater incentive to increase quality with attentionally differentiated competition. The second part of Proposition 10, that both types of competition generate higher quality than a monopoly, and therefore also higher quality than is optimal, is due to the effect of competition on prices. Since competition lowers prices, it shifts focus to quality, increasing firms’ incentives to produce higher quality.

More generally, this result says that in industries in which product differentiation is easier, there will be higher production of value because of the added gain from directing consumers’ attention toward one’s own strength. While we are not aware of a formal test of this prediction, it seems both intuitively reasonable and consistent with anecdotal examples. For instance, traditional telephones were difficult to differentiate, and the technology did not change for decades. However, following the introduction of mobile phones—which turned the phone into a personal item and opened a greater scope for differentiation—telephone technology improved rapidly.

²⁶ We have also explored the case when firms choose not only what quality level to produce, but also which attribute to produce on. In this case, there is no pure-strategy equilibrium. If firms choose the same dimension, then by Proposition 9 quality would be relatively low—but then switching to the other dimension and increasing quality is profitable. If firms produce on different dimensions, then quality would have to be high—but then switching to the other firm’s dimension and lowering quality is a profitable deviation.

6 A Model of Consideration Sets

In this section, we return to an important but so far underemphasized assumption of our model: that the decisionmaker’s focus and choices are based on an exogenously given consideration set that could be different from her choice set. Our motivation for allowing the consideration set to be different from the choice set is that manifestly irrelevant options in the consideration set have seemingly unrealistic effects on behavior. As a simple example, consider a choice set consisting of three options with two attributes each: $(0, 10)$, $(11, 0)$, and $(-1, -1000)$. If the consideration set was the entire choice set, the decisionmaker would put a very large weight on attribute 2, possibly leading her to choose the first option. It seems unlikely, however, that a person would focus so much on attribute 2 just because she has a terrible option on that attribute. For instance, a consumer who is looking to buy a new computer with a monitor will not focus more on the quality of monitors just because there are computers available without monitors. Hence, it seems more realistic to think of the decisionmaker’s consideration set as being made up of only the first two options.

More generally, in most economic settings a person has “self-destructive” choices that lower her utility in all respects, and these typically do not enter her consideration set—or often even her consciousness. While we feel that the approach of basing our model on an exogenously given consideration set consisting of the decisionmaker’s reasonable alternatives is satisfactory for many applications, for a complete theory of focus-driven choice it is necessary to define consideration sets endogenously. Because we are not aware of much evidence that would help guide such a formulation of consideration sets, we tentatively propose a theory as a first pass. But our theory above of how a person chooses from a consideration set can be combined with any theory of how the consideration set is formed.

The basic idea behind our formulation is simple, and builds on the premise of our theory that focus will be drawn to large differences. We amend this force by also assuming that a person’s attention is drawn to better rather than worse options. The consideration set, then, will be determined as the set that maximizes a combination of consumption utility and tradeoffs—that is, generates the most attractive tradeoffs.

Formally, suppose that the decisionmaker’s full choice set is \bar{C} , and let $\alpha \geq 1$. We normalize the consumption-utility-maximizing option in \bar{C} such that it has a consumption utility of zero. We assume that the consideration set is given by the set $C \subset \bar{C}$ that solves

$$\max_{C \subset \bar{C}} \frac{1}{|C|} \left[\alpha \sum_{c \in C} U(c) + \sum_{k=1}^K \Delta_k(C) \right]. \quad (3)$$

This definition has a number of potentially attractive properties. First, there always exists a consideration set, and generically, it is unique. Second, if $\alpha > 1$, c^1 dominates c^2 , and c^2 is in the consideration set, then so is c^1 ; otherwise, replacing c^2 by c^1 strictly increases the maximand in Equation 3. This also implies that if $\alpha > 1$, dominated choices can never constitute the consideration set. Third, if $\alpha = 1$ and there are two dimensions, the consideration set always includes the extreme options of the Pareto set.

Our theory of consideration-set determination is distantly related to that of Eliaz, Richter and Rubinstein (2010), who axiomatically characterize procedures for choosing a two-element consideration set (or “finalists”). Their “top two” criterion selects the best options according to some ordering, while their “extremes” criterion selects the two extremes according to some ordering. In trading off high utility and differences, our specification can be thought of as a combination of these two criteria.²⁷

With the endogenous theory of consideration sets, our model has a feature that makes it easier and methodologically less problematic to apply: once the function $g(\cdot)$ and the weight α are specified, applying our model to a new economic setting requires only the specification of relevant attributes in that setting in addition to assumptions (regarding the choice set and utility function) that a classical model would already make. Hence, our theory provides a way to translate a classical model into one with focus-sensitive choice with few degrees of freedom.

Context effects in purchasing. The following example illustrates the use of our consideration-set definition, and also identifies a new implication of our model. Consider a consumer who makes a purchase decision with five attributes, four product attributes and money. Suppose first that there are

²⁷ Different approaches include Eliaz and Spiegler (2011a) who assume that firms can influence consumers’ consideration set at a cost, and Hauser and Wernerfelt (1990) and Roberts and Lattin (1991) in the marketing literature, who analyze search-theory based models for consideration set formation.

two options, not buying and buying one particular product: $\bar{C} = \{(0, 0, 0, 0, 0), (20, 20, 0, 0, -35)\}$. For any $\alpha < 15$, the consideration set contains both options ($C = \bar{C}$). As a result, money looms large, and the consumer might not buy the product—even though buying it would maximize consumption utility. By a similar argument, the consumer might not buy if she has a choice between not buying and buying a more expensive product: $\bar{C} = \{(0, 0, 0, 0, 0), (0, 0, 22, 22, -40)\}$. But now suppose that *both* products are in her choice set: $\bar{C} = \{(0, 0, 0, 0, 0), (20, 20, 0, 0, -35), (0, 0, 22, 22, -40)\}$. In this case, for any $\alpha \geq 1$ the consideration set consists of only the last two options. Since money does not loom as large in this comparison, the decisionmaker might choose the more expensive option—even though in terms of consumption utility it is not worth the extra money.

Intuitively, when there are multiple products with sufficiently different attributes, the consumer’s attention is drawn to the tradeoffs between the products rather than the tradeoffs between buying and not buying. With the consumer considering *what* to buy rather than *whether* to buy, our model of focus implies that she will tend to care less about money, so that she is more likely to buy an expensive product. Both the prediction and the logic of this result are related to Bordalo’s (2011) salience-based explanation for the decoy effect, and Eliaz and Spiegel’s (2011b) theory of attention-grabbers.

7 Conclusion

By virtue of defining focus-dependent utility based on consumption utility and the decisionmaker’s choice set, our model opens the way for analyzing the role of focusing in many economic settings using one generally applicable model. In addition to the questions in intertemporal choice and product design we have started to address in this paper, there seem to be many additional potential applications. An employee may be motivated by some features of her employment contract—such as a bonus, major promotion, or other large goal—not only because they can generate higher consumption utility, but also because of her disproportionate focus on these features. A firm may recognize and take advantage of consumers’ distorted focus in specific situations, such as when selling an add-on to a consumer who is focusing too much on her current purchase. And just like marketers, political parties may attempt to manipulate voters by the positioning of their candidates

in part through attempting to attract focus to their candidates' strengths.

Being defined over riskless choices only, our model is not applicable to situations in which uncertainty is a central part. In ongoing work (Kőszegi and Szeidl 2011, hopefully), we develop an extension of our framework to uncertainty by thinking of states as different attributes, resulting in a model similar to that of Bordalo et al. (2010) and Bordalo (2011). Incorporating risk into our framework will allow us to apply it to a number of additional applications. For instance, because a large possible monetary loss in case of an accident attracts focus more than insurance payments in other states, a person may be too prone to buy insurance.

References

- Aaker, David A.**, *Managing Brand Equity*, New York: The Free Press, 1991.
- Acland, Dan and Matthew R. Levy**, "Habit Formation, Naivete, and Projection Bias in Gym Attendance," 2010. Working Paper, Harvard University.
- Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson**, "The Age of Reason: Financial Decisions Over the Lifecycle," 2007. MIT Department of Economics Working Paper No. 07-11.
- , — , — , and — , "Learning in the Credit Card Market," 2008. NBER Working Paper #13822.
- Anderson, James C., James A. Narus, and Wouter van Rossum**, "Customer Value Propositions in Business Markets," *Harvard Business Review*, March 2006, pp. 1–11.
- Asheim, Geir B.**, "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," 2008. Working Paper, University of Oslo.
- Bernheim, B. Douglas and Antonio Rangel**, "Beyond Revealed Preferences: Choice Theoretic Foundations for Behavioral Welfare Economics," January 2008. NBER Working Paper No. 13737.
- Bertaut, Carol C., Michael Haliassos, and Michael Reiter**, "Credit Card Debt Puzzles and Debt Revolvers for Self Control," *Review of Finance*, 2009, 13 (4), 657–692.
- Bordalo, Pedro**, "Choice-Set Effects and Saliency Theory," 2011. Working Paper, Harvard University.
- , **Nicola Gennaioli, and Andrei Shleifer**, "Saliency Theory of Choice Under Risk," 2010. Working Paper, Harvard University.

- Brown, Jeffrey R.**, “Understanding the Role of Annuities in Retirement Planning,” in Annamaria Lusardi, ed., *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs*, University of Chicago Press, 2009, chapter 6, pp. 178–208.
- , **Marcus D. Casey, and Olivia S. Mitchell**, “Who Values the Social Security Annuity? New Evidence on the Annuity Puzzle,” 2008. National Bureau of Economic Research Working Paper #13800.
- Bryan, Gharad, Dean Karlan, and Scott Nelson**, “Commitment Devices,” *Annual Review of Economics*, 2010, 2, 671–698.
- Chetty, Raj, Adam Looney, and Kory Kroft**, “Salience and Taxation: Theory and Evidence,” *American Economic Review*, 2009, 99 (4), 1145–1177.
- Chiu, Andrew and George Wu**, “Prospect Theory,” 2009. Entry for *Wiley Encyclopedia of Operations Research and Management Science*.
- Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond**, “Annuities and Individual Welfare,” *American Economic Review*, 2005, 95 (5), 1573–1590.
- DellaVigna, Stefano and Ulrike Malmendier**, “Contract Design and Self-Control: Theory and Evidence,” *Quarterly Journal of Economics*, 2004, 119 (2), 353–402.
- and —, “Paying Not to Go to the Gym,” *American Economic Review*, 2006, 96 (3), 694–719.
- Dunn, Elizabeth W., Timothy D. Wilson, and Daniel T. Gilbert**, “Location, Location, Location: The Misprediction of Satisfaction in Housing Lotteries,” *Personality and Social Psychology Bulletin*, 2003, 29, 1421–1432.
- Eliaz, Kfir and Ran Spiegler**, “Consideration Sets and Competitive Marketing,” *Review of Economic Studies*, 2011, 78 (1), 235–262.
- and —, “On the Strategic Use of Attention Grabbers,” *Theoretical Economics*, 2011, 6, 127–155.
- , **Michael Richter, and Ariel Rubinstein**, “Choosing the Two Finalists,” *Economic Theory*, 2010, 46 (2), 211–219.
- Federal Communication Commission**, “Comment Sought on Measures Designed to Assist U.S. Wireless Consumers to Avoid “Bill Shock”,” Technical Report 2010. Public Notice May 11 2010. CG Docket No. 09-158, http://hraunfoss.fcc.gov/edocs_public/attachmatch/DA-10-803A1.pdf.
- Gabaix, Xavier**, “A Sparsity-Based Model of Bounded Rationality,” 2011. Working paper, NYU Stern. URL: <http://pages.stern.nyu.edu/~xgabaix/papers/sparsebr.pdf>.

- , **David Laibson, Guillermo Moloche, and Stephen Weinberg**, “Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model,” *American Economic Review*, 2006, 96 (4), 1043–1068.
- Genesove, David and Chris Mayer**, “Loss Aversion and Seller Behavior: Evidence from the Housing Market,” *Quarterly Journal of Economics*, 2001, 116 (4), 1233–1260.
- Gourville, John T.**, “Pennies-a-Day: The Effect of Temporal Reframing on Transaction Evaluation,” *Journal of Consumer Research*, 1998, 24 (4), 395–408.
- Gruber, Jonathan and Botond Köszegi**, “Is Addiction ‘Rational?’ Theory and Evidence,” *Quarterly Journal of Economics*, 2001, 116 (4), 1261–1305.
- and — , “Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes,” *Journal of Public Economics*, 2004, 88 (9-10), 1959–1987.
- Hamermesh, Daniel S. and Joel B. Slemrod**, “The Economics of Workaholism: We Should Not Have Worked on This Paper,” *B.E. Journal of Economic Analysis and Policy: Contributions to Economic Analysis and Policy*, 2008, 8 (1).
- Hauser, John R. and Birger Wernerfelt**, “An Evaluation Cost Model of Consideration Sets,” *Journal of Consumer Research*, 1990, 16 (4), 393–408.
- Highfill, Jannett, Douglas Thorson, and William V. Weber**, “Tax Overwithholding as a Response to Uncertainty,” *Public Finance Review*, 1998, 26 (4), 376–391.
- Johnson, Eric J., John C. Hershey, Jacqueline Meszaros, and Howard Kunreuther**, “Framing, Probability Distortions, and Insurance Decisions,” *Journal of Risk and Uncertainty*, 1993, 7 (1), 35–51.
- Jones, Damon**, “Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds,” 2010. NBER Working Paper No. 15963.
- Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur A. Stone**, “Would You Be Happier If You Were Richer? A Focusing Illusion,” *Science*, 2006, 312 (5782), 1908–1910.
- and **Amos Tversky**, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 1979, 47 (2), 263–291.
- Köszegi, Botond and Ádám Szeidl**, “Attention and Compensation,” 2011, hopefully. In progress.
- Laibson, David I.**, “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 1997, 112 (2), 443–477.
- , **Andrea Repetto, and Jeremy Tobacman**, “Estimating Discount Functions with Consumption Choices over the Lifecycle,” 2007. NBER Working Paper #13314.

- Lehenkari, Mirjam**, “The Hedonic Editing Hypothesis: Evidence from the Finnish Stock Market,” *Journal of Behavioral Finance*, 2009, 10 (1), 9–18.
- Liberman, Nira and Yaacov Trope**, “The Role of Feasibility and Desirability Considerations in Near and Distant Future Decisions: A Test of Temporal Construal Theory,” *Journal of Personality and Social Psychology*, 1998, 75 (1), 5–18.
- Linville, Patricia W. and Gregory W. Fischer**, “Preferences for Separating or Combining Events,” *Journal of Personality and Social Psychology*, 1991, 60 (1), 5–23.
- Loewenstein, George, Ted O’Donoghue, and Matthew Rabin**, “Projection Bias in Predicting Future Utility,” *Quarterly Journal of Economics*, 2003, 118 (4), 81–123.
- Meier, Stephan and Charles Sprenger**, “Present-Biased Preferences and Credit Card Borrowing,” *American Economic Journal: Applied Economics*, 2010, 2 (1), 193–210.
- Odean, Terrance**, “Are Investors Reluctant to Realize Their Losses?,” *Journal of Finance*, 1998, 53 (5), 1775–1798.
- O’Donoghue, Ted and Matthew Rabin**, “Addiction and Self-Control,” in Jon Elster, ed., *Addiction: Entries and Exits*, Russel Sage Foundation, 1999.
- and — , “Optimal Sin Taxes,” *Journal of Public Economics*, 2006, 90 (10-11), 1825–1849.
- Rabin, Matthew and Georg Weizsäcker**, “Narrow Bracketing and Dominated Choices,” 2009. forthcoming, *American Economic Review*.
- Reis, Ricardo**, “Inattentive Producers,” *Review of Economic Studies*, 2006, 73 (3), 793–821.
- Roberts, John H. and James M. Lattin**, “Development and Testing of a Model of Consideration Set Composition,” *Journal of Marketing Research*, 1991, 28 (4), 429–440.
- Salant, Yuval**, “Procedural Analysis of Choice Rules with Applications to Bounded Rationality,” *American Economic Review*, forthcoming.
- Schkade, David A. and Daniel Kahneman**, “Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction,” *Psychological Science*, 1998, 9 (5), 340–346.
- Scitovsky, Tibor**, *The Joyless Economy: The Psychology of Human Satisfaction*, Oxford University Press, US, 1976.
- Shui, Haiyan and Lawrence M. Ausubel**, “Time Inconsistency in the Credit Card Market,” 2004. Available at <http://ssrn.com/abstract=586622>.
- Sims, Christopher A.**, “Rational Inattention and Monetary Economics,” in Benjamin A. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3A, Elsevier, 2010, chapter 4.

- Skiba, Paige Marta and Jeremy Tobacman**, “Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default,” 2008. Vanderbilt Law and Economics Research Paper No. 08-33.
- Spiegler, Ran**, *Bounded Rationality and Industrial Organization*, Oxford University Press, 2011.
- Stango, Victor and Jonathan Zinman**, “What Do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs,” *American Economic Review Papers and Proceedings*, 2009, *99* (2), 424–429.
- and —, “Fuzzy Math, Disclosure Regulation, and Market Outcomes: Evidence from Truth-in-Lending Reform,” *Review of Financial Studies*, 2011, *24* (2), 506–534.
- Thaler, Richard H.**, “Mental Accounting and Consumer Choice,” *Marketing Science*, 1985, *4* (3), 199–214.
- , “Psychology and Savings Policies,” *American Economic Review*, 1994, *84* (2), 186–192.
- , “Mental Accounting Matters,” *Journal of Behavioral Decision Making*, 1999, *12* (3), 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.
- Tversky, Amos and Daniel Kahneman**, “The Framing of Decisions and the Psychology of Choice,” *Science*, 1981, *211*, 453–458.
- Woodward, Susan E. and Robert E. Hall**, “Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence,” 2010. Working Paper, Stanford University.
- Yaari, Menahem E.**, “Uncertain Lifetime, Life Insurance, and the Theory of the Consumer,” *Review of Economic Studies*, 1965, *32* (5), 137–150.

A Proofs

Proof of Proposition 1. $1 \Rightarrow 2$. Throughout this proof, let $C = \{c^1, c^2\}$. The attention-weighted utility difference between c^1 and c^2 is

$$\begin{aligned} \tilde{U}(c^1, C) - \tilde{U}(c^2, C) &= \sum_{i \in A(c^1, c^2)} g(\Delta_i(C)) \cdot \Delta_i(C) - \sum_{j \in A(c^2, c^1)} g(\Delta_j(C)) \cdot \Delta_j(C) \\ &\geq g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{i \in A(c^1, c^2)} \Delta_i(C) - g\left(\max_{j \in A(c^2, c^1)} \Delta_j(C)\right) \cdot \sum_{j \in A(c^2, c^1)} \Delta_j(C) \\ &> g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{i \in A(c^1, c^2)} \Delta_i(C) - g\left(\min_{i \in A(c^1, c^2)} \Delta_i(C)\right) \cdot \sum_{j \in A(c^2, c^1)} \Delta_j(C) = 0, \end{aligned}$$

where in the strict inequality we use that $\min_{i \in A(c^1, c^2)} \Delta_i(C) > \max_{j \in A(c^2, c^1)} \Delta_j(C)$.

$2 \Rightarrow 1$. Suppose there are $\Delta \geq 0$ and $\Delta' > \Delta$ rational numbers such that $g(\Delta') \leq g(\Delta)$. Then, for a sufficiently large K we can construct a c^1 and c^2 with equal consumption utilities such $\Delta_i(C) = \Delta'(C)$ for $i \in A(c^1, c^2)$ and $\Delta_j(C) = \Delta(C)$ for $j \in A(c^2, c^1)$. This pair contradicts the statement of Part 2. By continuity, $g(\cdot)$ is everywhere strictly increasing. \square

Proof of Proposition 2. It is sufficient to prove that for any positive Δ and Δ' , we have

$$g(\Delta + \Delta')(\Delta + \Delta') > g(\Delta)\Delta + g(\Delta')\Delta'.$$

This is obvious since $g(\cdot)$ is increasing, so that $g(\Delta + \Delta') > g(\Delta), g(\Delta')$. \square

Proof of Proposition 3. Immediate. \square

Proof of Proposition 4. The attention-weighted utility difference between the two options is

$$\epsilon \cdot \sum_{k=1}^K g(\epsilon \cdot \delta_k) \cdot \delta_k.$$

Dividing by ϵ and letting $\epsilon \rightarrow 0$ gives

$$g(0) \cdot \sum_{k=1}^K \delta_k.$$

Since the sum of δ_k determines the consumption-utility-maximizing choice, we conclude that for ϵ small, the decisionmaker maximizes consumption utility. \square

Proof of Proposition 5.

1. Attention on dimension t is set by $\Delta_t = B$, while attention on dimensions $s = T_i + 1, \dots, T_i + T_b$ is determined by $\Delta_s = AB/T_b$. The consumer strictly prefers to invest if $g(B) < ABg(AB/T_b)$. Since the right-hand side is continuous and increasing in A , the value A_{post}^* is defined by $g(B) =$

$A^*B \cdot g(A^*B/T_b)$. It is immediate that A_{post}^* is independent of T_i . Given that $g(\cdot)$ is strictly increasing we also have that A_{post}^* is strictly increasing in T_b and equals 1 for $T_b = 1$.

2. Since all possible effort paths are considered, focus on the investment dimensions is set by $\Delta_t = B$, $t = 1, \dots, T_i$ while focus on the benefit dimensions is defined by $\Delta_s = AB \cdot T_i/T_b$ for $s = T_i + 1, \dots, T_i + T_b$. Effort is strictly preferred if and only if $g(B) < AB \cdot g(AB \cdot T_i/T_b)$, and A_{ante}^* is defined by $g(B) = A^*B \cdot g(A^*B \cdot T_i/T_b)$. The claims now all follow from the strict monotonicity of $g(\cdot)$.

3. Immediate from the above expressions. \square

Proof of Proposition 6. Consider $c'_t \in C_t$ such that $u_t(c'_t, \bar{c}_{t-1}^*) \leq u_t(c_t^*, \bar{c}_{t-1}^*)$. We show that in the ex-post decision of period t , the decisionmaker prefers c_t^* to c'_t . This in turn implies the statement of the proposition that the decisionmaker would prefer at least as high period- t utility ex post as she does ex ante.

Denote by \tilde{K} the set of attributes of period t utility which are affected by c_t . Both ex ante and ex post, the choice of c_t only affects period- t attributes in \tilde{K} and future attributes in periods $\tau > t$. The fact that ex ante the agent chooses c_t^* rather than c'_t means that

$$\sum_{k \in \tilde{K}} g(\Delta_{tk}) [u_{tk}(c_t^*, \bar{c}_{t-1}^*) - u_{tk}(c'_t, \bar{c}_{t-1}^*)] \geq \sum_{k, \tau > t} g(\Delta_{\tau k}) [u_{\tau k}(c'_t, \bar{c}_{\tau-t}^*) - u_{\tau k}(c_t^*, \bar{c}_{\tau-t}^*)].$$

By (ii), all the utility differences are non-negative on both sides. Now consider how the focusing weights change in this inequality in the ex-post choice. By (i), for each $k \in \tilde{K}$, the range of u_{tk} and hence the value of $g(\Delta_{tk})$ is the same in the ex-ante and the ex-post problem. In contrast, the values $\Delta_{\tau k}$ may be smaller ex post, when not all ex ante feasible consumption paths are considered. Thus the left hand side is unchanged while the right hand side is weakly smaller in the ex-post decision, and the inequality continues to hold. \square

Proof of Proposition 7. Consider $c'_t \in C_t$ such that $u_{\tau^*k^*}(c'_t) \leq u_{\tau^*k^*}(c_t^*)$. We show that in the ex-post decision of period t , the decisionmaker prefers c_t^* to c'_t . This in turn implies the statement of the proposition that the decisionmaker would prefer at least as high utility in attribute (τ^*, k^*) ex post as she does ex ante.

The fact that ex ante the agent chooses c_t^* rather than c'_t means that

$$g(\Delta_{\tau^*k^*}) [u_{\tau^*k^*}(c_t^*) - u_{\tau^*k^*}(c'_t)] \geq \sum_{(\tau, k) \neq (\tau^*, k^*), \tau > t} g(\Delta_{\tau k}) [u_{\tau k}(c'_t, \bar{c}_{\tau-t}^*) - u_{\tau k}(c_t^*, \bar{c}_{\tau-t}^*)].$$

By (ii) all terms are positive. By (i), the left-hand side is unchanged in the ex-post choice. Since the right-hand side is weakly smaller, the claim follows. \square

Proof of Proposition 8. The monopolist maximizes

$$\sum_{i=1}^{K_2} p_i - \phi \left(\sum_{i=1}^{K_1} v_i \right)$$

subject to

$$\sum_{i=1}^{K_2} G(p_i) \leq \sum_{i=1}^{K_1} G(v_i).$$

Denote $v = \sum_{i=1}^{K_1} v_i$. For any fixed v , Proposition 1 part 3 shows that the right-hand side of the constraint is maximized by setting $v_i = v$ for some i and $v_j = 0$ for $j \neq i$. The convexity of $G(\cdot)$ implies that for a fixed value of $\sum_{i=1}^{K_2} p_i$, the left hand side of the constraint is minimized when $p_i = p/j$ for all prices. It follows that the monopolist's problem is equivalent to maximizing

$$K_2 \cdot p - \phi(v)$$

subject to

$$G(v) = K_2 \cdot G(p).$$

Because $\phi'(0) = 0$, producing some value is better than producing no value. Because ϕ' is unbounded, it is not hard to see that producing a very large value is not optimal. It follows that this problem has an interior solution. Differentiability of the objective function and the constraint implies that the optimum must satisfy the first order condition.

Because $g(\cdot)$ is strictly increasing, the constraint implies that $v = p$ if $K_2 = 1$ and $v > p$ if $K_2 > 1$. Using the Lagrangian, the first order condition of this problem is

$$\phi'(v) = \frac{G'(v)}{G'(p)}.$$

This expression equates the marginal cost and marginal benefit of increasing value. The right hand side is the marginal benefit, because it measures by how much can prices be increased to keep the consumer indifferent, accounting for the differences in focus between the value and price attributes. For $K_2 = 1$ and $v = p$ this equation gives the efficient choice of value, while for $K_2 > 1$ and $v > p$ it implies overproduction since $G(\cdot)$ is strictly increasing. \square

Proof of Proposition 9. (i) By assumption, the focusing weights are set by \bar{v}_1 and \bar{v}_2 , and hence the consumer focuses more on attribute 1. Suppose that $v_1 < v_2$, and consider the deviation of increasing the first attribute and reducing the second by ϵ . This does not change the cost to firm 1, and if ϵ is small, also does not affect the focusing weights and hence the perceived value of firm 2's product. But this change does moves more of firm 1's value on the attribute which enjoys more attention, increasing the perceived value of the product.

(ii) Suppose by way of contradiction that $v_1 > v_2$. We will distinguish three cases. (a) Firm 1 is better only in the first attribute: $v_1 > \bar{v}_1$ but $v_2 \leq \bar{v}_2$. (b) Firm 1 is better only in the second attribute: $v_1 \leq \bar{v}_1$ and $v_2 > \bar{v}_2$. (c) Firm 1's product dominates: $v_1 > \bar{v}_1$ and $v_2 > \bar{v}_2$. We will only work out case (a) here; the other two cases are simpler. In all three cases, we consider the deviation where zero value is allocated to the first attribute and $v_1 + v_2$ is allocated to the second. This deviation has the same cost as before, and it is straightforward to check that it is profitable in cases (b) and (c) as long as $v_1 + v_2 > \bar{v}_1 + \bar{v}_2$. Consider case (a). The perceived value difference between the two firms' products before the deviation is

$$g(v_1)(v_1 - \bar{v}_1) + g(\bar{v}_2)(v_2 - \bar{v}_2)$$

while after the deviation it equals

$$g(\bar{v}_1)(-\bar{v}_1) + g(v_1 + v_2)(v_1 + v_2 - \bar{v}_2).$$

After manipulations, the difference can be written as

$$(\bar{v}_1 - \bar{v}_2)(g(v_1) - g(\bar{v}_1)) + (v_1 - \bar{v}_2)(g(v_1 + v_2) - g(v_1)) + v_2(g(v_1 + v_2) - g(\bar{v}_2)) - \bar{v}_2(g(\bar{v}_1) - g(\bar{v}_2)).$$

The first three terms are positive and the last is negative. When \bar{v}_2 goes to zero, the last term vanishes and hence the deviation is profitable. When $g(v_1 + v_2)$ increases without bound, either the first or the second term becomes unbounded, while the last term remains constant, and hence the deviation is profitable in that case as well. \square

Proof of Proposition 10. Consider a candidate symmetric equilibrium (v, p) for attentionally undifferentiated competition. Due to competition, we must have $p = \phi(v)$. We solve for conditions under which firm 1 would not want to deviate from this candidate equilibrium. Take first a deviation to $v^1 > v$ and $p^1 > p$. For the consumer to be indifferent between the two products, we must have

$$g(v^1)v^1 - g(p^1)p^1 = g(v^1)v - g(p^1)p,$$

so that

$$g(v^1)(v^1 - v) - g(p^1)(p^1 - p) = 0.$$

Differentiating totally with respect to v^1 and evaluating at $v^1 = v, p^1 = p$ gives

$$\left. \frac{dp^1}{dv^1} \right|_{v^1=v} = \frac{g(v)}{g(p)}.$$

For (v, p) to be an equilibrium, it must be the case that by this local deviation, the firm cannot make a positive profit; that is, $\phi'(v) \geq g(v)/g(p)$.

A similar calculation for a local deviation to $v^1 < v, p^1 < p$ shows that we must also have $\phi'(v) \leq g(v)/g(p)$, so that equilibrium requires

$$\phi'(v) = \frac{g(v)}{g(\phi(v))}. \tag{4}$$

Now consider attentionally differentiated competition. Consider again a candidate equilibrium (v, p) with $p = \phi(v)$. If firm 1 deviates to $v^1 > v, p^1 > p$, then for the consumer to be indifferent between the two products we must have

$$g(v^1)v^1 - g(p^1)p^1 = g(v)v - g(p)p.$$

Unlike in the case of attentionally undifferentiated competition, here a deviation by firm 1 to increase value *decreases* rather than increases attention on the competitor's product. Rewriting the above expression gives

$$g(v^1)v^1 - g(p^1)(p^1 - p) = g(v)v.$$

Differentiating totally with respect to v^1 and evaluating at $v^1 = v, p^1 = p$ gives

$$\left. \frac{dp^1}{dv^1} \right|_{v^1=v} = \frac{g(v) + g'(v)v}{g(p)},$$

and the same expression holds for downward deviations. As a result, an equilibrium must satisfy

$$\phi'(v) = \frac{g(v) + g'(v)v}{g(\phi(v))}. \quad (5)$$

The statements of the proposition follow from comparing (4) and (5) with the analogous expression in the proof of Proposition 8. \square

B Eliciting Model Ingredients from Behavior

Since our model has ingredients that are not based directly on observed behavior, in this section we outline an algorithm for eliciting these ingredients from choices. Once the ingredients are elicited, our model provides a prediction on behavior for any finite choice set. Note that from the perspective of falsifiability, this is logically equivalent to providing axiomatic foundations for the model.

The intuitive idea behind our elicitation is to notice that because the agent makes consumption-utility-maximizing choices in balanced decisions, we can elicit consumption utility from behavior much like in a standard setting. Once we know the consumption utilities, we can identify the focus weights from choice behavior by eliciting how the willingness to pay for changes in an attribute depends on the preexisting difference in that attribute across options. To see how a consumer's monetary valuation for a one-attribute product can be elicited, for instance, consider the choice between utility vectors $(u, -v(p))$ and $(0, 0)$, where the former option is buying the product with consumption utility u at price p and the latter option is not buying. This is a balanced choice with $K' = 1$, $a = u$ and $b = v(p)$. Hence, at the price for which the consumer is indifferent between buying and not buying, we must have $u = v(p)$, so that p is the monetary value of the product based on consumption utility.

Formally, for our elicitation we assume that there are at least three attributes. Our elicitation requires that we know how products map into attributes, and that we can manipulate individual attributes in the decisionmaker's choices. We also assume that the function $g(\cdot)$ is strictly increasing and it and the utility functions u_k are twice differentiable, although the elicitation works somewhat more generally with minor modifications. We normalize $u_k(0) = 0$ for each k , $g'(0) = 1$, and $u_1'(0) = 1$.

The first step in our algorithm elicits the curvature of the utility function for each dimension k . Focusing only on dimensions 1 and k , consider choice sets of the form $\{(0, x + \delta(p)), (p, x)\}$ for any $x \in \mathbb{R}$ and $p > 0$. For any $p > 0$ we can find the $\delta(p)$ that makes the decisionmaker indifferent between the two options. Hence, we have

$$g(u_1(p) - u_1(0))(u_1(p) - u_1(0)) = g(u_k(x + \delta(p)) - u_k(x))(u_k(x + \delta(p)) - u_k(x)),$$

which implies

$$u_1(p) - u_1(0) = u_k(x + \delta(p)) - u_k(x)$$

since $g(\cdot)$ is strictly increasing. Dividing the above by p and letting $p \rightarrow 0$ gives

$$u'_1(0) = u'_k(x)\delta'(p).$$

Hence, this procedure elicits $u'_k(x)$, and hence (using the normalization that $u_k(0) = 0$) the entire utility function $u_k(\cdot)$.

The second step in our procedure elicits the attention weights $g(\cdot)$. Since we have now elicited the utility function, here we work directly with utilities. Looking only at dimensions 1, 2, and 3, consider choice sets of the form $\{(p, 0, x), (0, x + \delta(p), 0)\}$. Again, for each p we find the $\delta(p)$ that makes the decisionmaker indifferent between the two options. Hence

$$G(p) = G(x + \delta(p)) - G(x),$$

where $G(y) \equiv g(y)y$. Differentiating,

$$G'(p) = G'(x + \delta(p))\delta'(p),$$

and differentiating again,

$$G''(p) = G''(x + \delta(p)) (\delta'(p))^2 + G'(x + \delta(p))\delta''(p).$$

Setting $p = 0$ gives

$$2 = 2g'(0) = G''(0) = G''(x)(\delta'(0))^2 + G'(x)\delta''(0),$$

which identifies $G(\cdot)$ and hence also $g(\cdot)$.

To conclude, we note that even if the attributes of some options are not fully known, we may be able to use knowledge about the known attributes of other options to also extract information about the unknown attributes. As an example that is highly relevant for some of our applications in intertemporal choice, suppose that an employee will receive a \$5,000 bonus, which she will spend over following 10 months in equal installments. Suppose further that we do not know whether the employee thinks of the bonus at least in part as a single one-time payment or as a 10-time increase in consumption of \$500. We can use the bias toward concentration to elicit this from behavior. Specifically, we can elicit her willingness to work for each additional \$500 in consumption in the 10 months. If her willingness to work for the \$5,000 bonus is the same as her total willingness to work for the increases in consumption, she thinks of the bonus merely in terms of the consumption. But if her willingness to work for the bonus is greater than her total willingness to work for the increases in consumption, she thinks of the payment in part as a single attribute.²⁸

²⁸ Some assumptions in our model, however, cannot be elicited from choice behavior, and therefore must come from outside our theory. For instance, even if we know that a scholar derives utility from completing a book in part because she views it as a “book achievement” attribute on which no other option provides utility, we cannot tell whether in her mind there are one or two such book achievement attributes. Psychologically, however, it seems that a person would view such essentially identical attributes as one, so that there is only a single book achievement attribute.

C Naive Dynamic Model

In this section, we formulate a theory of future beliefs that is an alternative to the rational beliefs we have posited in our model of focusing in intertemporal choices in Section 2.1. In this new specification, the individual holds incorrect that correspond to an ex-ante view that derives from the commitment solution to her problem. Intuitively, a naive agent formulates her beliefs about future behavior at some ex-ante stage, when she has a general global view of her decision problem. Hence, her beliefs are based on the ex-ante commitment problem. But when she makes a specific choice in a specific period, her decision is based on a local view of her decision problem. Although different in the specific theory of behavior, this perspective of the decisionmaker's thinking is reminiscent of construal level theory in psychology as applied to intertemporal choice. Liberman and Trope (1998), for instance, argue that temporally distant events are construed by individuals at an abstract, broad level, whereas nearby events are construed on more specific terms.

Formally, we assume that the decisionmaker's beliefs derive from the ex-ante commitment problem, in which focus weights are generated by the set C_{ante}^* of possible lifetime consumption profiles. But unlike when solving for the single ex-ante optimal consumption profile, to have a full mapping from current behavior to future behavior we must now specify behavior at each date for each past contingency. This is straightforward. For any vector of choices c_1, \dots, c_t up to period t , let $\tilde{C}_{t+1}(c_t, \dots, c_1)$ be the feasible future consumption paths. Then, for any vector of past choices c_1, \dots, c_{t-1} , we require the beliefs $(c_{t+1}(c_t), \dots, c_T(c_t))$ to solve

$$\max_{c'_{t+1}, \dots, c'_T \in \tilde{C}(c_1, \dots, c_t)} \tilde{U}((c_1, \dots, c_t, c'_{t+1}, \dots, c'_T), C_{ante}^*).$$